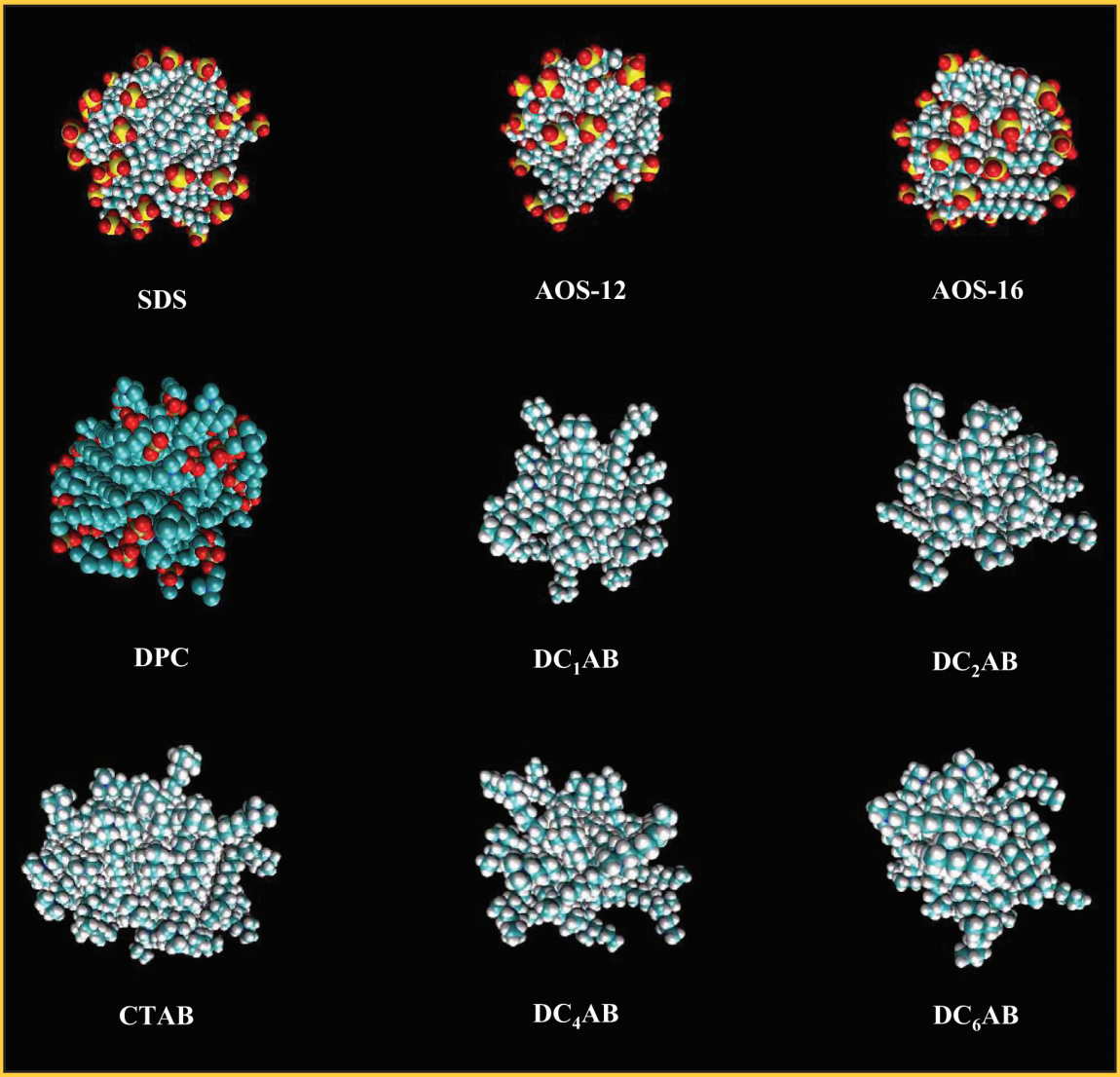
JPCBFK

VOLUME 111
FEBRUARY 8, 2007
NUMBER 5

http://pubs.acs.org/JPCB

THE JOURNAL OF PHYSICAL CHEMISTRY B



Post-Equilibration Structures of the Surfactant Micelles Simulated to Implement the CS–MT Model (see page 1063)

CONDENSED MATTER AND BIOPHYSICAL CHEMISTRY

PUBLISHED WEEKLY BY THE AMERICAN CHEMICAL SOCIETY

Quantifying the Hydrophobic Effect. 1. A Computer Simulation−Molecular-Thermodynamic Model for the Self-Assembly of Hydrophobic and Amphiphilic Solutes in Aqueous Solution

Brian C. Stephenson, Arthur Goldsipe, Kenneth J. Beers, and Daniel Blankschtein*

Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received: September 1, 2006; In Final Form: November 21, 2006

Surfactant micellization and micellar solubilization in aqueous solution can be modeled using a molecular-thermodynamic (MT) theoretical approach; however, the implementation of MT theory requires an accurate identification of the portions of solutes (surfactants and solubilizates) that are hydrated and unhydrated in the micellar state. For simple solutes, such identification is comparatively straightforward using simple rules of thumb or group-contribution methods, but for more complex solutes, the hydration states in the micellar environment are unclear. Recently, a hybrid method was reported by these authors in which hydrated and unhydrated states are identified by atomistic simulation, with the resulting information being used to make MT predictions of micellization and micellar solubilization behavior. Although this hybrid method improves the accuracy of the MT approach for complex solutes with a minimum of computational expense, the limitation remains that individual atoms are modeled as being in only one of two states—head or tail—whereas in reality, there is a continuous spectrum of hydration states between these two limits. In the case of hydrophobic or amphiphilic solutes possessing more complex chemical structures, a new modeling approach is needed to (i) obtain quantitative information about changes in hydration that occur upon aggregate formation, (ii) quantify the hydrophobic driving force for self-assembly, and (iii) make predictions of micellization and micellar solubilization behavior. This article is the first in a series of articles introducing a new computer simulation−molecular thermodynamic (CS−MT) model that accomplishes objectives (i)−(iii) and enables prediction of micellization and micellar solubilization behaviors, which are infeasible to model directly using atomistic simulation. In this article (article 1 of the series), the CS−MT model is introduced and implemented to model simple oil aggregates of various shapes and sizes, and its predictions are compared to those of the traditional MT model. The CS−MT model is formulated to allow the prediction of the free-energy change associated with aggregate formation (g_{form}) of solute aggregates of any shape and size by performing only two computer simulations—one of the solute in bulk water and the other of the solute in an aggregate of arbitrary shape and size. For the 15 oil systems modeled in this article, the average discrepancy between the predictions of the CS−MT model and those of the traditional MT model for g_{form} is only 1.04%. In article 2, the CS−MT modeling approach is implemented to predict the micellization behavior of nonionic surfactants; in article 3, it is used to predict the micellization behavior of ionic and zwitterionic surfactants.

1. Introduction

Surfactants are molecules consisting of a hydrophilic moiety, referred to as the head, attached to a hydrophobic moiety, referred to as the tail. This dual nature of surfactants leads to very interesting behavior in aqueous solution. Above a threshold surfactant concentration, known as the critical micelle concentration, or CMC, the surfactant molecules self-assemble into aggregates known as micelles with their hydrophobic tails partly shielded from water in the aggregate interior (the aggregate core) and their hydrophilic heads exposed to water at the aggregate surface.

The solubility of chemicals that have limited solubility in aqueous solution can be increased through the addition of surfactants.[1−4] These chemicals are frequently referred to as solubilizates, and they may be either completely hydrophobic or amphiphilic (containing both hydrophilic and hydrophobic moieties). Amphiphilic solubilizates may behave much like conventional surfactants within a micellar environment, and their incorporation into micelles can be accurately modeled using theories developed originally to model mixed surfactant micellization.[5] However, unlike surfactants, solubilizates have a solubility limit, rather than a CMC, in aqueous solution, and without added surfactant, they do not spontaneously self-assemble to form micelles.

Gaining a fundamental understanding of the process of micellization and micellar solubilization in aqueous solution is of both academic and practical interest. The most predictive and accurate theoretical models of surfactant micellization and micellar solubilization implement what is known as the molecular-thermodynamic (MT) modeling approach.[8,14] In the MT modeling approach, the free-energy change associated with the formation of the surfactant aggregate is expressed as the sum of several free-energy contributions, all of which can be computed molecularly given the chemical structures of the various micellar components. MT theory enables prediction of

* Corresponding author. Department of Chemical Engineering, Room 66-444, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139. Telephone: (617) 253-4594. Fax: (617) 252-1651. E-mail: dblank@mit.edu

10.1021/jp065696i CCC: $37.00
Published on Web 01/18/2007

equilibrium properties of self-assembled surfactant/solubilizate systems in aqueous solution on the basis of the chemical structures of the solution components and the solution conditions (such as the temperature, pressure, and ionic strength).[8−13] Among the equilibrium micellar solution properties that can be predicted are the CMC, the micelle size distribution, the micelle shape and average size, the extent of micellar solubilization, and the locus (or location) of solubilization within a micelle. The MT model introduced by Nagarajan and Ruckenstein permits modeling of micellar aggregates composed of nonionic, zwitterionic, and ionic surfactants.[8] In recent years, our group has also contributed to the development of MT models to predict surfactant behavior in aqueous solution.[9,14−22] Theoretical efforts have been most successful at modeling the self-assembly of relatively simple surfactant systems, including surfactants that have linear hydrocarbon (or fluorocarbon) tails and a single, rigid head, for example, sodium dodecyl sulfate (SDS) and cetyltrimethylammonium bromide (CTAB). Some progress has also been made in modeling surfactants with long polymeric heads, including those of the alkyl poly(ethylene oxide) variety.[8]

An alternative to MT descriptions of micellization is the application of computer simulation approaches to simulate the self-assembly of surfactants and solubilizates into micelles. In theory, molecular dynamics and Monte Carlo computer simulations based on an atomistic force field have the advantage of being capable of modeling arbitrarily complex chemical structures and enabling quantitatively or semiquantitatively accurate predictions of micellization and solubilization phenomena. However, although simulations of micelle self-assembly with atomistic-level detail are possible, such simulations have only been performed well above the CMC because simulation time is severely limited by the size and density of micellar systems.[46] To accurately identify the CMC, extended simulations of large surfactant/solvent systems would be necessary at a range of low surfactant concentrations and over prohibitively long time scales.[46] Consequently, prediction of surfactant solution properties directly through computer simulations of surfactant self-assembly is not feasible at the present time.

The hydrophobic effect, or the increase in solution free energy observed upon addition of nonpolar solutes to water, is the primary driving force responsible for surfactant self-assembly in aqueous solution.[6] It is also the primary driving force responsible for solubilizate incorporation into surfactant micelles in aqueous solution. In addition to the hydrophobic effect, the process of micelle self-assembly is also mediated by van der Waals, hydrogen-bonding, and screened electrostatic interactions (in the case of charged surfactants).[7]

Because the hydrophobic effect is the primary driving force for micelle self-assembly in aqueous solution, it is essential to accurately model this contribution to the overall free energy of micelle formation in any theoretical or computer simulation description of micelle formation. In order to model the hydrophobic effect in the context of the MT approach, a reasonable a priori determination must be made about the way in which the surfactant and the solubilizate molecules are hydrated in the micellar state. By comparing the degree of hydration of various groups within each solute in the micellar state with the degree of hydration of those same groups in the bulk aqueous solution, the changes in hydration that occur upon micelle self-assembly can be determined and MT theory can be used to quantify the hydrophobic driving force for micelle formation.

To date, MT models of micellization and micellar solubilization have relied on relatively simple approximations for the micellar hydration states of the surfactants and the solubilizates. In the traditional MT approach, each surfactant molecule is modeled as being composed of two distinct portions—the head and the tail. The surfactant head is considered to be fully hydrated in both the monomeric and the micellar states. The surfactant tail is considered to be at least partially dehydrated in the micellar state, with the degree of dehydration being a function of the micelle geometry.[8,14] To identify the head and the tail, one approach involves determining the relative degree of hydrophobicity of different groups within a solute molecule using a group-contribution approach such as the one included in the software package Molecular Modeling Pro.[23] Using this type of information, one can make educated guesses about which portions of a simple solute are hydrated in the micellar state. For example, for surfactants with an alkyl group attached to a charged or zwitterionic head, group-contribution approaches suggest that the first CH_2 group attached to the charged or dipolar head also possesses hydrophilic character, and the remainder of the CH_2 groups and the terminal CH_3 group remain hydrophobic. On the basis of this information, as well as on the basis of some experimental evidence for charged and zwitterionic surfactants,[5,14] the approximation is made that $n_t = n_c - 1$, where n_c is the total number of CH_2 and CH_3 groups in the hydrocarbon chain and n_t is the number of CH_2 and CH_3 groups that should be modeled as being part of the surfactant tail. For surfactants with an alkyl group attached to a nonionic head, every CH_2 and CH_3 group is assumed to be part of the tail, such that $n_t = n_c$.[22]

Unfortunately, in the case of more complex solute chemical structures, making head and tail assignments using simple group-contribution methods is inadequate. Examples of surfactants and solubilizates for which making head and tail assignments is not trivial are shown in Figure 1. In the case of the surfactant alkyl-3-hydroxy sulfonate (AOS), the presence of the two hydrophobic CH_2 groups between the two hydrophilic groups (SO_3^- and OH) makes the head and tail identification challenging. In the case of the surfactant decanoyl-*n*-methylglucamide (MEGA-10), it is difficult to determine the micellar hydration state of each of the three groups bonded to the nitrogen atom (CH_3, CH_2, and the carbonyl group). In the case of the solubilizates *o*-, *m*-, and *p*-aminobenzoate, it is unclear what effect changing the relative locations of the NH_2 group and the ethyl ester group within the molecule has on the head and tail identification. In addition, it is unclear whether the ethyl ester group attached to the benzene ring should be modeled as being part of the solubilizate head or tail.

To extend the applicability of the MT modeling approach to more chemically and structurally complex surfactants and solubilizates, there is a need to accurately estimate the hydration states of these solutes in the micellar state. Even for the relatively simple surfactants and solubilizates shown in Figure 1, the prediction of such hydration information is beyond the scope of simple group-contribution methods, because the hydration states of the various chemical groups in the micelle are intimately related to the connectivity of these groups within a given solute. Although it may be possible to develop a suitable group-contribution approach that accounts for this connectivity in order to predict the required hydration information, such an approach must be parametrized based on a training set of detailed micellar hydration data for relatively complex surfactants and solubilizates. Unfortunately, at the present time, such data is not available. Fortunately, however, atomistic-level computer simulations provide a promising approach to gather such information.

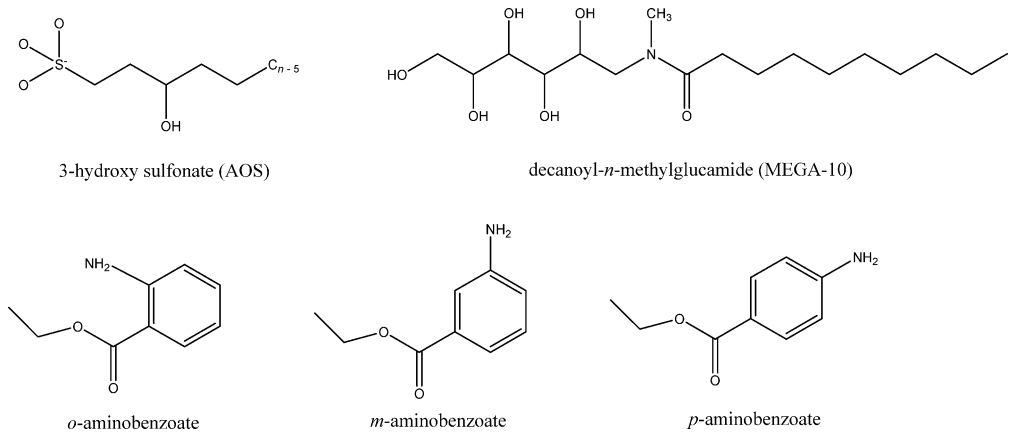


Figure 1. Examples of surfactants and solubilizates for which making head and tail assignments is not trivial. As a result, the surfactants shown here are difficult to model using the traditional MT modeling approach.

With the above need in mind, we recently reported the development of a traditional MT modeling approach in which molecular dynamics (MD) computer simulations of surfactant molecules at an oil−water interface (serving as a proxy for the hydrophobic micelle core−water interface) were used to determine the head and the tail groups of both structurally simple and relatively complex surfactants.[24] Subsequently, this approach was also used to determine the head and the tail groups of the pharmaceutically relevant solubilizate ibuprofen.[25] In each case, traditional MT modeling was conducted on the basis of computer simulation assignments of heads and tails, and the theoretical modeling results were compared with the experimental data.[24,25] Reasonable agreement was obtained between the theoretical predictions and the experimental data, with our results indicating that accurate head and tail assignments for traditional MT modeling can indeed be made through the use of MD simulations.[24,25]

To overcome the limitations of MT modeling (namely, its applicability to relatively simple surfactant and solubilizate systems) and of computer simulations (namely the high computational cost of modeling self-assembly at atomistic detail), in this article, we present a novel theoretical approach that combines hydration information determined through the use of computer simulation (CS) with a new molecular-thermodynamic (MT) model used to quantify the hydrophobic effect. This theoretical approach, which will be referred to hereafter as the CS−MT model, can be used to predict the self-assembly behavior of solutes in aqueous solution and is presented generally in this article for hydrophobic and amphiphilic solutes, including oil molecules, surfactants, and solubilizates.

The CS−MT model extends previous MT modeling work in the following important ways:

(i) By using hydration data obtained from computer simulation, the CS−MT modeling approach avoids making some of the hydration approximations that have been made in traditional MT models of micellization and micellar solubilization. In addition, the computer simulation results allow one to evaluate the accuracy of the hydration approximations made in traditional MT modeling.[14]

(ii) A new theoretical model is developed to make quantitatively accurate estimates of the hydrophobic driving force involved in the self-assembly of hydrophobic and amphiphilic solutes in aqueous solution using hydration data obtained from computer simulation. The CS−MT model decomposes the contribution to the free energy from the hydrophobic effect into two parts: (a) a "dehydration" free-energy change associated with the decrease in the number of solute−water contacts with respect to the dilute aqueous solution limit and (b) a hydration free-energy change describing the change in the free energy of the remaining solute−water contacts as compared to those in a dilute aqueous solution. Using this modeling approach, we can model solute groups that are partially hydrated and reside the majority of the time at the micelle core−water interface in a more physically realistic manner than is currently possible using the traditional MT model.

(iii) Through computer simulation of the micellar state, a wealth of information about micelle microstructure can be obtained and used as an input to refine traditional MT or CS−MT modeling. Examples of such structural information are the projected area of each solute head at the micelle core−water interface, the distribution of conformations of the solute tails within the micelle core, the degree of counterion binding (in the case of ionic solutes), and the locus of solubilization. Although the effect of using such structural information to improve MT modeling is not explored in this article, we are pursuing such improvements as part of our ongoing research.

To validate and test the implementation of the CS−MT model, we will present CS−MT and traditional MT modeling results for 15 different oil aggregates of various shapes (spheres, cylinders, and slabs) and sizes. These oil aggregates were selected as a starting point to validate the CS−MT modeling approach because they are significantly simpler to model than micelles, in which the presence of the surfactant heads introduces additional complications. In article 2 of this series,[26] we will describe the implementation of the CS−MT model to predict the micellization behavior of simple and relatively complex nonionic surfactants (where electrostatic effects are absent) in aqueous solution. Finally, in article 3 of this series, we will describe the implementation of the CS−MT model to predict the micellization behavior of ionic and zwitterionic surfactants (where electrostatic effects are present).[27]

The remainder of the article is organized as follows. The traditional MT modeling approach for solute self-assembly is reviewed in Section 2, including a description of the thermodynamic framework underlying the traditional MT and the CS−MT models (Section 2.2) and an overview of the traditional MT model (Section 2.3). The CS−MT model is introduced in Section 3. The model is formulated generally for a wide class of solutes, including oil molecules, nonionic and ionic surfactants, and nonionic and ionic solubilizates. Section 4 describes the computer simulation approach used to obtain the hydration information required in the CS−MT model, including an overview of the modeling approach (Section 4.2), the simulation methods and parameters (Section 4.3), and a description of how each system was

prepared and equilibrated (Section 4.4). The data analysis method used to analyze the MD trajectories is described in Section 4.5. Computer simulation results are presented in Section 5. In Section 6, the CS−MT model is used to model oil aggregates, results are compared with those obtained using the traditional MT model, and the validity and accuracy of the CS−MT model is discussed. Concluding remarks are presented in Section 7. Finally, the CS−MT model is extended to allow the prediction of aggregate shape and size in Appendix A, and the equivalence of the CS−MT and the traditional MT modeling approaches is demonstrated mathematically in the case of completely hydrophobic solutes in Appendix B.

2. Molecular-Thermodynamic Modeling Approach

2.1. Introduction. In this section, we review the traditional MT model, with particular emphasis on how the hydrophobic effect is quantified. The MT model presented here is applicable to a broad class of hydrophobic and amphiphilic solutes, including oil molecules, surfactants, and solubilizates, although as discussed in Section 1, it can only be applied to relatively simple solutes without additional information about the hydration states of each solute in the micellar environment.

The central objective of the MT modeling of micellization and solubilization in aqueous solution is to compute g_{form}, the free-energy change associated with transferring the surfactant monomers, the solubilizates, and any bound counterions (in the case of ionic surfactants) from their standard states in the aqueous solution to a micellar aggregate in its standard state.[5,14,22] Quantification of the hydrophobic contribution to g_{form} is essential to accurately model this transfer process. The MT modeling approach relies on a thermodynamic framework to describe the micellar solution.[14,15,22] This thermodynamic framework allows the calculation of useful micellar solution properties, including the critical micelle concentration (CMC), the distribution of aggregate shapes and sizes, and microstructural characteristics of the micellar aggregate (such as the locus of solubilization) from g_{form}. A brief overview of the thermodynamic framework is presented in Section 2.2. Note that this framework is formulated in the context of modeling a binary mixture of a single surfactant species and a single solubilizate species but can be reformulated in a straightforward manner to model single surfactant micellization, the micellization of *n*-component surfactant mixtures, and the micellar solubilization of *n*-component mixtures of surfactants and solubilizates.[5,14,22] After introducing the thermodynamic framework, we briefly review the traditional MT modeling approach in Section 2.3.

2.2. Thermodynamic Framework. In the multiple-chemical equilibrium model of micellization,[9,14] each micellar aggregate is considered to be a distinct chemical species in equilibrium with the other aggregates and the individually dispersed solutes present in the aqueous solution. By equating the chemical potentials of the micellar aggregates, the surfactant monomers (s), the solubilizate monomers (sol), and the counterions (c), an expression is obtained that describes the mole fraction of micellar aggregates, $X_{n_s\alpha\beta}$, containing n_s surfactant molecules, n_{sol} solubilizate molecules, and βn_s bound counterions (where $\alpha = n_s/(n_{sol} + n_s)$ is the aggregate composition and β is the degree of counterion binding). Specifically,[5]

$$X_{n_s\alpha\beta} = \left(\frac{1}{e}\right) X_{1s}^{n_s} \exp\left[-\frac{n_s g_{form}(S, l_c, \alpha, \beta)}{k_B T}\right] \quad (1)$$

where k_B is the Boltzmann constant, T is the absolute temperature, and g_{form} is defined as follows:

$$g_{form} = \left[\frac{\mu^o_{n_s\alpha\beta}}{n_s} - \mu^o_s - \beta\mu^o_c - \frac{\alpha}{1-\alpha}\mu^o_{sol}\right] - k_B T - \beta k_B T \ln(X_{1c} e) - \frac{\alpha}{1-\alpha} k_B T (X_{1sol} e) \quad (2)$$

In eq 2, μ_i^o is the standard-state chemical potential of species i (where i refers to a $n_s\alpha\beta$ micellar aggregate, a surfactant monomer, a solubilizate monomer, or an unbound counterion). The variables X_{1s}, X_{1sol}, and X_{1c} in eqs 1 and 2 are the mole fractions of the surfactant monomers, the solubilizate monomers, and the counterions, respectively. As shown in eq 2, g_{form} is a function of the aggregate shape (S), the aggregate core-minor radius (l_c), the aggregate composition (α), and the degree of counterion binding (β).

At the values of S, l_c, α, and β that minimize g_{form} (denoted as S^*, l_c^*, α^*, and β^*), g_{form} has an optimal value denoted hereafter as g^*_{form}. Due to the exponential dependence of $X_{n_s\alpha\beta}$ on ($n_s g_{form}$) in eq 1, small deviations from g^*_{form} yield $X_{n_s\alpha\beta}$ values that are essentially zero. Accordingly, by solving for g^*_{form}, the optimal aggregate shape (S^*), the optimal core-minor radius (l_c^*), the optimal composition (α^*), and the optimal degree of counterion binding (β^*) can be predicted. In addition, the CMC in mole fraction units is computed as follows:[22]

$$CMC \approx \exp\left(\frac{g^*_{form}(S^*, l_c^*, \alpha^*, \beta^*)}{k_B T}\right) \quad (3)$$

2.3. Traditional MT Model of Surfactant Micellization and Micellar Solubilization. MT theory can be used to predict g_{form} based on the chemical structures of each of the solutes in the aqueous solution—whether they are oil molecules, surfactants, or solubilizates. As discussed in Section 1, important inputs to the MT model are the hydrated and unhydrated portions of each solute in the micellar state.[24] In the traditional MT modeling approach, g_{form} is expressed as the sum of the following six free-energy contributions:[22]

$$g_{form} = g_{tr} + g_{int} + g_{pack} + g_{st} + g_{elec} + g_{ent} \quad (4)$$

Each of the six contributions in eq 4 arises from a distinct step in a thermodynamic cycle used to model the process of micelle formation. The various steps involved are shown schematically in Figure 2, which depicts the micelle formation process for a binary mixture of a cationic surfactant and a nonionic hydrophobic solubilizate in aqueous solution. An analogous thought process may be used to model the formation of a multicomponent surfactant micelle or a multicomponent surfactant/solubilizate micelle.

In the first step shown in Figure 2, the cationic surfactant heads are separated from the surfactant tails and subsequently discharged along with the negative counterions in the aqueous solution. The corresponding discharge free energy is denoted as $g_{discharge}$.[15,16]

In the second step shown in Figure 2, a hydrophobic micelle core composed of the surfactant and solubilizate tails (referred to collectively as the solute tails) is formed. This step is modeled as the sum of three free-energy contributions: g_{tr}, g_{int}, and g_{pack}. The transfer free-energy contribution, g_{tr}, represents the free-energy change associated with transferring the solute tails from the aqueous solution to a bulk solution of solute tails.[6] The interfacial free-energy contribution, g_{int}, represents the free-energy change associated with forming an interface between the solute tails and the aqueous solution.[14] The packing free-

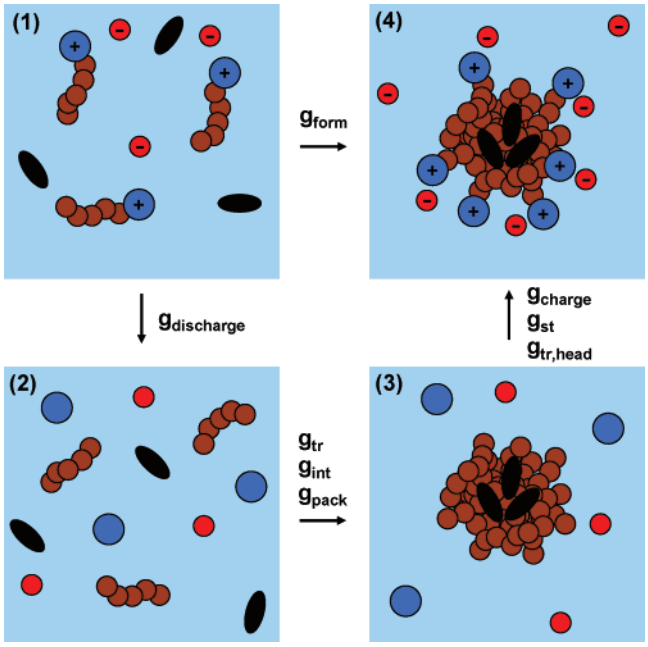


Figure 2. Sequence of steps followed in the MT cycle used in the CS−MT modeling approach developed in this article. This sequence is presented in the context of the micellization of a cationic surfactant and a nonionic hydrophobic solubilizate. Between frames 1 and 2, the solute heads (the large blue circles carrying positive charges) are separated from the solute tails (the chains consisting of five brown circles), and the solute heads and the counterions (the small red circles carrying negative charges) are discharged (as reflected in $g_{discharge}$). Between frames 2 and 3, the solute hydrophobic tails and the hydrophobic solubilizates (the black ovals) are grouped to form the micelle core (as reflected in g_{tr}, g_{int}, and g_{pack}). Between frames 3 and 4, the solute heads are reattached to one end of the solute tails (as reflected in g_{st} and $g_{tr,head}$), and the solute heads and their associated counterions are recharged (as reflected in g_{charge}).

energy contribution, g_{pack}, represents the free-energy change required to fix one end of the amphiphilic solute tails (in the example considered here, only the surfactant tails) at the micelle core−water interface. This free-energy contribution is estimated using a mean-field model first introduced by Ben-Shaul, Szleifer, and Gelbart[29−31] and requires sampling each important conformation and orientation of the solute tail subject to the constraint that the hydrophobic micelle core has uniform density.

In the third step shown in Figure 2, the surfactant heads are transferred to the surface of the micelle (with a corresponding free-energy contribution, g_{st}) and recharged along with the counterions (with a corresponding free-energy contribution, g_{charge}).[15,16,32] The steric free-energy contribution, g_{st}, accounts for the steric penalty associated with placing the surfactant heads in close proximity at the aggregate core−water interface.[33] In localizing the surfactant heads at the aggregate core−water interface, the heads are transferred to a different environment than that corresponding to the bulk water reference state. The change in free energy associated with this transfer corresponds to $g_{tr,head}$. However, in the traditional MT modeling approach, the surfactant heads are assumed to remain fully hydrated in the aggregate state, and therefore, $g_{tr,head}$ is approximated as being equal to zero. As a result, the free-energy contribution, $g_{tr,head}$, is not listed in eq 4. We define the electrostatic free-energy contribution, g_{elec}, in eq 4 as being equal to the sum of $g_{discharge}$ and g_{charge}.[15,16] Note that the entropic free-energy contribution, g_{ent}, although included in eq 4, is not shown in Figure 2 because it can contribute to the thermodynamic cycle at several stages. The entropic free-energy contribution includes the translational entropy loss incurred by the solubilizates upon association with the micelles, the translational entropy loss of the bound counterions, and the mixing entropy associated with a multicomponent micelle.[15,16] A more detailed description of the conceptual thought process implemented in the traditional MT modeling approach can be found in refs 14 and 15.

It is important to note that the salient characteristic of a solubilizate in the context of the MT model is that it has a solubility limit in aqueous solution rather than a CMC. Chemically, solubilizate molecules can be amphiphilic like conventional surfactant molecules. If an amphiphilic solubilizate is present in the micelle, it may localize at the micelle core−water interface in a manner analogous to a surfactant molecule. If this is the case, only a portion of the amphiphilic solubilizate molecule will be transferred to the hydrophobic micelle core and contribute to g_{tr} and $g_{int.}$ In addition, because one end of the solubilizate tail is constrained to reside at the micelle core−water interface, the solubilizate tails also contribute to g_{pack}. Finally, the solubilizate head contributes to g_{st} and $g_{tr,head}$, and its charge state contributes to g_{elec}.

The sequence of steps outlined above have been used by our group for many years to describe the process of micelle formation.[9,14−22] Recently, Maibaum et al. used a similar sequence of steps to model micelle formation, although the thermodynamic cycle that they proposed was used solely to model nonionic surfactants.[34] Specifically, Maibaum et al. broke up the formation of the hydrophobic micelle core into two steps: the formation of a vapor cavity in aqueous solution and the transfer of the solute tails to this vapor cavity. The free energy of formation of the vapor cavity was modeled as being equal to the water−vapor surface tension times the surface area of the vapor cavity, plus a pressure−volume work term associated with cavity formation (which is negligible for water at standard conditions). After filling the hydrophobic core with the solute tails, the water−vapor surface tension was corrected to become equal to the interfacial tension of a water−oil interface.

An important aspect of the thermodynamic cycle shown in Figure 2 that we would like to highlight is that, in conceptually separating the solute heads from the tails in aqueous solution, and in subsequently reattaching them in the micellar environment, we do not allow the hydration states of the solute tails to change. Therefore, the hydration states of the solute tails are assumed to be the same in frames 1 and 2 as well as in frames 3 and 4 in Figure 2. The transition involved in moving from frame 2 to frame 3 in Figure 2 reflects the formation of the micelle hydrophobic core. The changes in hydration incurred in the formation of this micelle core represent the primary driving force for micelle formation. In traditional MT modeling, this driving force is modeled by the two terms, g_{tr} and g_{int}. Because a more general and accurate calculation of the hydrophobic contribution to the free energy of micelle formation is the central aim of this article, we will discuss the traditional MT modeling approach used to calculate g_{tr} and g_{int} in more detail in Sections 2.3.1 and 2.3.2, respectively. For a detailed discussion of the other free-energy contributions appearing in eq 4, the interested reader is referred to refs 14, 15, and 22.

2.3.1. The Transfer Free-Energy Contribution, g_{tr}. In the traditional MT modeling approach, only the solute tails contribute to g_{tr}. Therefore, to determine this free-energy contribution, it is first necessary to identify the head and tail of each component present in the micelle. Various approaches for such identification were described in Section 1. After identifying the solute tails, g_{tr} is estimated as a composition-weighted average of g_{tr} for each micellar component expressed on a per surfactant molecule basis. For example, in a micelle containing a single

surfactant type and a single solubilizate type, the transfer free-energy contribution is estimated as follows:[5]

$$g_{tr} = g_{tr,s} + \frac{(1-\alpha)}{\alpha} g_{tr,sol} \quad (5)$$

where $g_{tr,s}$ and $g_{tr,sol}$ are the free-energy contributions associated with transferring a surfactant (s) tail and a solubilizate (sol) tail, respectively, to the aggregate core. Note that we divided by α in eq 5 to obtain an expression on a per surfactant molecule basis. For linear alkyl tails, correlations have been developed to express solubility as a function of alkyl chain length, temperature, and the concentration of added salt in aqueous solution.[14,35,36] For more complex solutes, more sophisticated group-contribution methods or experimental data may be used to estimate the tail solubility.[24,25] Note that aqueous solubility is related to g_{tr} by the relationship $g_{tr} = k_B T \ln(s)$, where s is the aqueous solubility of the solute expressed on a mole fraction basis.

2.3.2. The Interfacial Free-Energy Contribution, g_{int}. Some, or all, of the solute tails that are transferred to the aggregate core will reside for some time at the micelle core−water interface. The free-energy penalty associated with partially rehydrating these tail moieties is referred to as the interfacial free-energy contribution (g_{int}) and is modeled in the traditional MT approach using a micelle core−water interfacial tension. In a two-component micelle containing a single surfactant type and a single solubilizate type, g_{int} is computed on a per surfactant molecule basis using the following expression:[5]

$$g_{int} = (a - a_0)\left[\frac{\alpha^{int}\sigma_s + (1-\alpha^{int})\sigma_{sol}}{\alpha}\right] \quad (6)$$

where a is the area available to each solute at the micelle core−water interface, a_0 is the interfacial area that is screened by the solute heads on a per surfactant molecule basis, α^{int} is the mole fraction of surfactant at the micelle core−water interface, and σ_j is the curvature-dependent interfacial tension between water and a bulk phase of solute tails of type j (where j = s or sol).[14] Note that we have divided by α in eq 6 to obtain an expression on a per surfactant molecule basis. The curvature-dependent interfacial tension, σ_j, is determined using the Gibbs−Tolman−Koenig−Buff equation:[37−40]

$$\sigma_j = \frac{\sigma_{0,j}}{\left(1 + \frac{(S-1)\delta}{l_c}\right)} \quad (7)$$

where $\sigma_{0,j}$ is the interfacial tension of component j at a flat interface (having a typical value of about 50 mN/m for hydrocarbons), δ is the Tolman distance,[40] and S is a shape factor (3 for spheres, 2 for cylinders, and 1 for disks or bilayers). We typically use an empirical correlation to determine $\sigma_{0,j}$ for alkyl chains of varying length and as as a function of temperature, although, if available, the experimental $\sigma_{0,j}$ values may be used.[41] The Tolman distance, δ, is computed using the following expression:[14]

$$\delta(n_t) = \delta(n_t = 11) l_{max}(n_t)/l_{max}(n_t = 11) \quad (8)$$

where n_t is the number of carbons in the solute tail and $l_{max}(n_t) = 1.54 + 1.265\, n_t$ is the fully extended length of the solute tail (in Å).[6]

3. Computer Simulation−Molecular Thermodynamic (CS−MT) Modeling Approach

3.1. Introduction. In this section, we describe a new model for the hydrophobic effect that incorporates hydration information obtained from computer simulation as an input. This new model provides a more general approach to quantify the hydrophobic driving force for aggregate self-assembly than what is traditionally used in MT modeling, which was reviewed in Section 2. We develop this new model using a general expression for the extent of hydration. After introducing the new modeling approach in this Section, the simulation approach used to obtain the extent of hydration data required to implement the CS−MT model is discussed in Section 4.

3.2. Theoretical Framework. As discussed in Section 2.3 and shown in eq 4, in the traditional MT modeling approach, the hydrophobic effect is quantified by the free-energy contributions, g_{tr} and g_{int}. In order to more accurately quantify the hydrophobic effect, we propose to replace the terms g_{tr}, $g_{tr,head}$, and g_{int} (see Figure 2) with (i) the free-energy contribution associated with dehydration (g_{dehydr}) and (ii) the free-energy contribution associated with hydration (g_{hydr}), that is, to rewrite eq 4 as follows:

$$g_{form} = g_{dehydr} + g_{hydr} + g_{pack} + g_{st} + g_{elec} + g_{ent} \quad (9)$$

Because we propose to use computer simulation to determine g_{dehydr} and g_{hydr}, we refer to this modeling approach as the CS−MT model. The models used to calculate g_{dehydr} and g_{hydr} are presented in Sections 3.3 and 3.4, respectively. Each of the remaining four free-energy contributions in eq 9 are identical to those included in the traditional MT model for g_{form}. Nevertheless, it is important to stress that, in principle, computer simulations could also be used to improve the estimation of g_{pack} and g_{st} (see below).

As discussed in Section 2.3, g_{pack} is the free-energy penalty associated with fixing one end of the solute tails at the aggregate core−water interface.[15] In the traditional MT modeling approach, each conformation and orientation of the solute tails (subject to the constraint of constant aggregate core density) is sampled after fixing one end of each tail at the aggregate core−water interface, and the free-energy difference between the constrained and the unconstrained states is estimated using a mean-field description.[29−31] An alternative to generating solute tail conformations in a separate computational step and using the generated conformations to determine g_{pack} would be to use the tail conformations recorded during an MD simulation to compute g_{pack}. Such an approach could potentially allow computation of g_{pack} with less computational expense than is currently possible and would also allow a more seamless integration of the computer simulation and the MT modeling approaches. Although this approach has not yet been implemented, we are investigating the utility of incorporating this type of approach as part of our ongoing modeling work.

The steric free-energy contribution, g_{st}, accounts for the free-energy penalty associated with placing the solute heads in close proximity at the aggregate core−water interface. To determine this contribution, the cross-sectional area of the solute head at the aggregate core−water interface must be estimated. In the case of small, rigid heads, it is straightforward to do so from knowledge of the head chemical structure. However, in the case of larger, polymeric heads, such estimation is quite challenging. In such cases, the head area can be estimated from the computer simulation results by calculating the average projected area of each solute head at the aggregate core−water interface over

the course of the molecular dynamics simulation run. We are investigating this use of computer simulations as part of our ongoing modeling work.

Computer simulation results of the structure of the aggregate state can also be used to check key structural predictions made by traditional MT theory, including (i) the degree of counterion binding (when ionic solutes are present in the micelle) and (ii) the locus of solubilization (when solubilizates are present in the micelle). By comparing the MD and the MT predictions, it is possible to ascertain to what extent both approaches yield consistent and physically reasonable results. Accurate MT estimates of the degree of counterion binding and of the locus of solubilization are essential to accurately estimate the free-energy contributions in eq 9. In future publications involving ionic surfactants and solubilizates, we will compare the structural predictions of the MD and the MT modeling approaches and discuss the implications of our results on both the CS−MT and the traditional MT modeling approaches.

3.3. The Free Energy of Dehydration, g_{dehydr}. The dehydration free-energy contribution in eq 9 is necessary to account for the change in free energy associated with the dehydration that accompanies self-assembly into aggregates. In the CS−MT model, we quantify hydration using the following general expression, which we refer to as the fractional hydration, f, where

$$f = \frac{\text{extent of hydration in the aggregate environment}}{\text{extent of hydration in the bulk water environment}} \quad (10)$$

Note that $f = 0$ corresponds to complete dehydration and $f = 1$ corresponds to the extent of hydration in the bulk water environment. The fractional dehydration in the aggregate environment is equal to $(1 - f)$. The specific manner in which f will be estimated using computer simulation data will be discussed in detail in Section 4. However, using the definition of the fractional hydration given in eq 10, we present here a theoretical model to allow estimation of the hydrophobic driving force for aggregate formation.

Every hydrophobic group in a solute tail and most groups in a solute head will experience some degree of dehydration upon aggregate formation. As we discussed in Section 2.3, in traditional MT theory, the heads are assumed to remain fully hydrated in the micellar state and do not contribute to the hydrophobic driving force for aggregate formation (in other words, $g_{tr,head} = 0$ in Figure 2). Using the hydration information obtained through computer simulation data, we find that it is no longer necessary to make this approximation in the context of the CS−MT modeling approach. Indeed, here, we propose a more general approach to quantify the free-energy contribution associated with the dehydration of *any* hydrophobic group in the solute, regardless of whether the group is in the solute head or in the solute tail. Specifically, we propose the following model for g_{dehydr}:

$$g_{dehydr} = \sum_{i=1}^{n_{hydr}} (1 - f_i) g_{tr_i} \quad (11)$$

where n_{hydr} is the total number of hydrophobic groups in the solute, $(1 - f_i)$ is the fractional dehydration associated with group i upon aggregate formation, and g_{tr_i} is the free-energy change associated with transferring group i from the aqueous solution to a bulk phase of hydrophobic tails. Note that, in eq 11, hydrophilic groups are assumed to have a negligible effect on the dehydration free energy and, therefore, are not included in the summation. The validity of this approximation will be discussed in greater detail in article 2 of this series.

To compute g_{dehydr} using eq 11, it is necessary to accurately estimate g_{tr_i} for each hydrophobic group in the solute molecule. For the oil molecules considered in this article, g_{tr_i} is only needed for the CH_2 and the CH_3 groups that comprise the oil molecule, and is obtained using the same solubility correlations for linear alkyl tails that are used in the traditional MT modeling approach.[35]

3.4. The Free Energy of Hydration, g_{hydr}. The hydration free-energy contribution in eq 9 is necessary to account for the difference in free energy associated with hydrating contacts experienced in bulk water and with hydrating contacts experienced after incorporation into the core of an aggregate. Hydrating contacts have a different free energy in the two states because the size of a single hydrophobic chain in water is much smaller than the size of an aggregate core. This difference in size, in turn, induces a different extent of disruption of the surrounding water molecules in the two states, leading to a different hydration free energy. The size dependence of hydration thermodynamics is a well-known phenomenon and has been modeled theoretically in an approach developed by Lum, Chandler, and Weeks (the LCW Theory).[42] For small solutes (typically smaller than 1 nm in radius), the solute volume is sufficiently small that it does not disrupt the hydrogen-bonding network in the surrounding water molecules. For larger solutes (or clusters of solutes), the hydrophobic surface is of sufficiently low curvature that it disrupts the hydrogen-bonding network, reducing the density of water near the surface and creating a solute−water interface.[43]

We propose the following model to compute g_{hydr}:

$$g_{hydr} = \sum_{i=1}^{n_{core}} SASA_i f_i \Delta g_{wc_i} \quad (12)$$

where n_{core} is the total number of hydrophobic groups in the solute that adsorb onto, or are incorporated into, the aggregate core; $SASA_i$ is the solvent accessible surface area of group i; f_i is the fractional hydration associated with group i upon aggregate formation; and Δg_{wc_i} is defined as the difference in the free energy per unit of solvent accessible surface area associated with hydration in the aggregate state and in the aqueous solution for group i.

To the extent that a hydrophobic group adsorbed onto, or incorporated into, the aggregate core remains hydrated in the aggregate state, the free-energy change associated with hydrating contacts for that group is accounted for with the term Δg_{wc_i}. For amphiphilic solutes (which contain both a head and a tail), only those hydrophobic groups that are actually incorporated into the aggregate core contribute to g_{hydr}. Any hydrophobic groups that extend away from the aggregate core into the aqueous solution are best modeled as having the same free energy associated with water contacts in the aggregate state as in the bulk aqueous state, because the extent of disruption of the hydrogen-bonding−coordinate-bonding network of the solution in both states is very similar.

If desired, the term Δg_{wc_i} in eq 12 may be used as a fitting parameter to obtain the closest possible agreement between (i) the traditional MT model and the CS−MT model or (ii) the experimental data and the theoretical predictions of the CS−MT model. However, in order to develop a predictive model that requires no experimental input, we propose a theoretical approach to estimate Δg_{wc_i} for the oil molecules considered in this article.

In general, Δg_{wc_i} depends on the chemical nature of group i. However, the only molecules of interest in this article are oils. Because of the chemical similarity of the CH_2 and the CH_3 groups in oil, Δg_{wc_i} can be approximated as being equal for both groups. Note that the size difference between the CH_2 and the CH_3 groups is accounted for through the $SASA_i$ term in eq 12. With this approximation in mind, in the remainder of this article, we will refer to Δg_{wc_i} as Δg_{wc} when modeling each oil aggregate considered.

By predicting Δg_{wc} theoretically for oil aggregates and evaluating the accuracy of the CS−MT modeling results, we will be able to assess the validity and range of applicability of eqs 9, 11, and 12. In addition, we will be able to assess the validity of the computer simulation approach that will be described in Sections 4 and 5. We propose the following theoretical model for Δg_{wc}:

$$\Delta g_{wc} = \sigma_{core} - \sigma_{bulk} = \frac{\sigma A_{core}}{SASA_{core}} + \frac{g_{tr_i}}{SASA_i} \qquad (13)$$

where σ_{core} is the "microscopic interfacial tension" (interfacial free energy per unit SASA) associated with the aggregate hydrophobic core−water interface, σ_{bulk} is the microscopic interfacial tension (interfacial free energy per unit SASA) associated with the group i (CH_2 or CH_3)−water interface in the aqueous solution, σ is the macroscopic interfacial tension of the aggregate hydrophobic core−water interface (computed using eq 7), A_{core} is the area of the aggregate hydrophobic core computed geometrically on the basis of the volume of the aggregate subject to the assumption of a perfectly smooth aggregate surface, $SASA_{core}$ is the solvent accessible surface area of the aggregate hydrophobic core, and $SASA_i$ is the solvent accessible surface area of group i.

In eq 13, $\sigma_{core} = \sigma A_{core}/SASA_{core}$ and $\sigma_{bulk} = -g_{tr_i}/SASA_i$ (recall that g_{tr_i} is < 0). By defining $\sigma_{core} = \sigma A_{core}/SASA_{core}$, we demand that the microscopic interfacial tension experienced at the aggregate interface be equal to the free energy of the aggregate hydrophobic core−water interface per unit of interfacial SASA. By defining $\sigma_{bulk} = -g_{tr_i}/SASA_i$, we demand that the microscopic interfacial tension experienced in bulk water at the group i−water interface be equal to the negative of the transfer free energy of group i per unit of solute SASA. The difference between σ_{core} and σ_{bulk} is equal to the free-energy difference per unit SASA associated with the hydrating contacts in the aggregate state and in the bulk water state.

The validity of eq 13 hinges on whether it is physically reasonable to evaluate Δg_{wc} on a per unit SASA basis, thereby invoking the concept of a microscopic interfacial tension, or microscopic interfacial free energy per unit area. Modeling the aggregate hydrophobic core as having a microscopic interfacial tension (σ_{core}) is reasonable given the size of the oil aggregates considered here (1.15 to 1.48 nm in radius), the typical size of hydrophobic micellar cores (~1 nm in radius), and the success obtained in modeling the micellar hydrophobic core using a curvature-corrected interfacial tension in the traditional MT modeling approach (see eq 7).[14] However, because a linear alkane is only ~0.25 nm in radius in its smallest dimension,[14] it is more questionable to model the solvation free energy of an oil molecule in the bulk aqueous solution as being proportional to SASA. For very small solutes, past research suggests that the solvation free energy can be modeled more accurately as being linearly related to the solvated volume rather than to the solvated surface area.[43] Nevertheless, a number of researchers, including Tanford, have modeled solubility as a function of SASA for linear and branched alkyl chains with reasonable accuracy, suggesting that the relatively simple model proposed in eq 13 is adequate.[44,45] The validity of eq 13 will be discussed in greater detail in Section 5 on the basis of the values of σ_{core} and σ_{bulk} determined from our computer simulation results. Although σ_{bulk} is approximately constant, given the dependence of σ on the alkyl tail length and on the aggregate hydrophobic core curvature, σ_{core} is also expected to be a function of alkyl tail length and curvature. In addition, we note that the ratio of $A_{core}/SASA_{core}$ may also be a function of these variables. In Section 5, we will present results for $A_{core}/SASA_{core}$ for 15 oil aggregates of different shapes (spheres, cylinders, and slabs) and sizes. Using the computer simulation results, we will then be able to estimate Δg_{wc} using eq 13.

3.5. Extension of the CS−MT Model to Predict Aggregate Shape and Size. It is important to note that the theoretical framework that we have presented above allows one to determine g_{form} only for an aggregate for which hydration data is available from computer simulation. Because of the computational expense associated with performing atomistic computer simulations of aggregate systems, it is not practical to perform simulations of many aggregates having different shapes and sizes to identify the aggregate geometry that corresponds to the minimum value of g_{form}, which in turn, corresponds to the aggregate geometry that will be realized experimentally.

A salient capability of traditional MT theory is that it enables prediction of g_{form} as a function of aggregate shape and size.[14,15] From this known functional dependence, it is then possible to predict the optimal aggregate shape and size. In Appendix A, we outline a computational strategy to extend the CS−MT modeling approach to enable prediction of g_{form} as a function of aggregate shape and size. This is accomplished by combining elements of the CS−MT and the traditional MT models.

3.6. Evaluating the Validity and Accuracy of the CS−MT Model. To validate and test the implementation of the CS−MT model, we have selected a total of 15 different oil aggregates for simulation and modeling. Oil aggregates were selected as a starting point to validate the CS−MT modeling approach because of their simplicity. Indeed, unlike nonionic or ionic surfactant micelles, oil aggregates are devoid of solute heads at the aggregate core−water interface. As a result, for each of the 15 oil aggregates considered, there are only two contributions to g_{form}: g_{tr} and g_{int} in the traditional MT modeling approach (see eq 4) and g_{dehydr} and g_{hydr} in the CS−MT modeling approach (see eq 9). In other words, g_{pack}, g_{st}, g_{elec}, and g_{ent} in eqs 4 and 9 are all equal to zero, because the oil aggregates are (i) devoid of solute heads at the aggregate core−water interface (g_{pack} and $g_{st} = 0$), (ii) nonionic ($g_{elec} = 0$), and (iii) single-component systems ($g_{ent} = 0$).

Unfortunately, experimental data for g_{form} is not available for these oil aggregates because these structures are not thermodynamically stable except at infinite dilution. As a result, in this article, we will compare the predictions of the CS−MT model with those of the traditional MT model. However, in article 2 of this series, we will discuss the implementation of the CS−MT model to predict the micellization behavior of nonionic surfactants. In that case, the CS−MT model predictions will be compared with both the traditional MT model predictions as well as with experimental CMC data. CMC predictions were selected as the micellar property of interest in article 2 because the CMC depends exponentially on g_{form} (see eq 3), and as such, it provides an excellent quantitative metric with which to assess the predictive accuracy of the CS−MT model.

The 15 different oil aggregates that were selected for modeling include a total of five geometries: (i) a spherical oil

aggregate of radius ≈ 1.16 nm (the "small" spherical oil aggregate), (ii) a spherical oil aggregate of radius ≈ 1.45 nm (the "large" spherical oil aggregate), (iii) a cylindrical oil aggregate of radius ≈ 1.10 nm (the "small" cylindrical oil aggregate), (iv) a cylindrical oil aggregate of radius ≈ 1.30 nm (the "large" cylindrical oil aggregate), and (v) a planar slab of oil with a half-width ≈ 0.85 nm (the "planar" oil aggregate). In addition, we selected three different types of oil molecules for simulation and modeling in the five chosen aggregate geometries: octane, dodecane, and hexadecane. The three oil molecules selected cover a range of alkyl chain lengths that are frequently encountered in surfactant and solubilizate tails.

By conducting simulations for three different types of oil molecules in five different geometries possessing different curvatures, we will be able to thoroughly evaluate the accuracy of the CS−MT modeling approach for the types of hydrocarbon tails and aggregate geometries that are most commonly encountered in modeling micellization and micellar solubilization. In addition, by modeling spherical oil aggregates, cylindrical oil aggregates, and planar oil slabs, we will be able to evaluate the ability of the CS−MT approach to model the three idealized micellar geometries (a perfect sphere, a perfect cylinder, or a perfect bilayer) that are used in the context of the traditional MT modeling approach.[9,14,15,22]

4. Molecular Dynamics Simulations

4.1. Introduction. In this section, we describe a general molecular dynamics simulation approach that can be used for any hydrophobic or amphiphilic solute to determine the detailed hydration information required to more accurately quantify the hydrophobic driving force responsible for self-assembly. This approach is introduced here in the context of the 15 different oil aggregates that were selected to validate the CS−MT model (see Section 3.6). The approach presented here will also be used to obtain hydration information for nonionic surfactants in article 2 of this series.

4.2. Modeling Approach. To more accurately quantify the hydrophobic driving force associated with the formation of each of the 15 aggregates considered here, we have used atomistic computer simulations to determine the change in hydration of each atom (or group of atoms, such as a CH_2 group) that is transferred from the aqueous solution environment to the aggregate environment. To accomplish this, two simulations were performed. The first one—the "bulk water" simulation—was of a single solute in a simulation cell containing water. The second one—the "aggregate" simulation—was of the same solute in the aggregate environment. Because it is very computationally expensive to simulate the self-assembly of solutes in aqueous solution into aggregates, for each of the aggregates simulated here, the aggregate was preformed in vacuum and subsequently equilibrated in a box of water.[46] Although these equilibration times were sufficient to allow rearrangement and equilibration of the solutes within the aggregate, they were not sufficient to allow a solute to leave the aggregate environment and enter the aqueous environment. Accordingly, the simulated aggregate will not necessarily have the same geometry (shape and size) that would be observed experimentally.

4.3. Simulation Methods and Parameters. All the solutes considered were modeled using bonded and nonbonded interaction potentials included in the fully atomistic OPLS-AA force field.[47] Water was modeled using the simple extended point-charge (SPC/E) model for water. SPC/E represents an improvement over SPC in which a correction is implemented to account for the self-polarization of water.[48] Atomic charges were assigned to each oil molecule on the basis of the default atomic charge values recommended in OPLS-AA. Van der Waals interactions were incorporated using a cutoff distance of 0.9 nm, and Coulombic interactions were described using the 3D particle mesh Ewald (PME) summation.[49,50] Although the van der Waals cutoff used in this study is shorter than the cutoffs that are frequently reported in the literature (1.2 to 1.4 nm), there is a tradeoff between using longer cutoff distances to try to more accurately capture the nonbonded interactions present in the system and using the same cutoff distances used for the original force field parametrization. It has been demonstrated that truncation schemes for electrostatic interactions give qualitatively incorrect results when compared with newer and more accurate methods, such as reaction field treatment of electrostatics or Ewald summation.[51] However, using relatively short-range cutoffs for van der Waals interactions yields accurate results with the inclusion of long-range dispersion corrections, as shown in recent simulation studies using the OPLS-AA force field carried out by Shirts et al.[52−54] In our simulations, long-range dispersion corrections were implemented to more accurately model the energy and the pressure of the system. Both dispersion corrections are negative, and although the energy correction is small, the pressure correction is significant and must be included to yield accurate results.[55] In modeling short-range, nonbonded interactions, a neighbor list with a cutoff of 0.9 nm was maintained and updated every 10 simulation steps. Each simulation was carried out with fixed bond lengths using the SHAKE algorithm as implemented in GROMACS,[56] which allowed an increase in simulation time step from 1 to 2 fs.

In each simulation, the cell temperature was maintained at 298.15 K using a Berendsen temperature coupling algorithm, which mimics weak coupling to an external heat bath with first-order kinetics.[55] A Berendsen pressure coupling algorithm was used to maintain each simulation cell at the desired pressure of 1.0 bar.[55] All simulations were conducted using a 2006 developers' version of the GROMACS software package.[57,58]

4.4. System Preparation and Equilibration. *4.4.1. Bulk Water Simulations.* The "bulk water" simulation for each of the oil molecules considered was initialized by placing a single oil molecule in a simulation cell and surrounding it with water molecules. The simulation cell was selected to be sufficiently large that there would always be at least 2.0 nm separating the oil molecule from its periodic image. Computer simulation studies of the propagation of water ordering away from an interface suggest that such a separation distance should be sufficient to prevent the oil molecule from interacting with its periodic image.[59] After a brief equilibration under *NPT* conditions until the system volume had stabilized, a 2 to 5 ns data-gathering simulation was conducted.

4.4.2. Aggregate Simulations. The geometric characteristics of the five simulated aggregates of octane, dodecane, and hexadecane are listed in Tables 1, 2, and 3, respectively. In total, 15 different aggregate systems were prepared and simulated. It is important to note that, to the extent possible (given the requirement of an integer number of molecules), aggregation numbers for octane, dodecane, and hexadecane were selected such that corresponding aggregate geometries would have the same dimensions. In other words, the small spherical octane aggregate has the same dimensions as the small spherical hexadecane aggregate. This was done to permit direct comparison of the simulation results for the different oil molecules considered and to evaluate the effect of the hydrocarbon chain length on these results. The volume, V, of each aggregate reported in Tables 1, 2, and 3 was computed using the formula

TABLE 1: Overview of Octane MD Simulations[a]

aggregate type	n	V [nm^3]	A [nm^2]	l_c [nm]	PL [nm]	C [nm^{-1}]
small sphere	25	6.75	17.27	1.17		1.71
large sphere	50	13.5	27.41	1.48		1.35
small cylinder	54	14.58	26.57	1.10	3.85	0.91
large cylinder	54	14.58	22.83	1.28	2.85	0.78
slab	192	51.84	28.88	0.90	3.80	0.0

[a] Geometric characteristics of each of the five octane aggregates simulated, including the aggregation number (n), the aggregate volume (V), the aggregate surface area (A), the aggregate core-minor radius (l_c), the periodic length (PL) of each cylinder and slab, and the degree of curvature (C, as defined in eq 14).

TABLE 2: Overview of Dodecane MD Simulations[a]

aggregate type	n	V [nm^3]	A [nm^2]	l_c [nm]	PL [nm]	C [nm^{-1}]
small sphere	17	6.42	16.70	1.15		1.73
large sphere	33	12.46	25.99	1.44		1.39
small cylinder	36	13.59	25.66	1.06	3.85	0.94
large cylinder	36	13.59	22.04	1.23	2.85	0.81
slab	128	48.33	28.88	0.84	3.80	0.0

[a] Geometric characteristics of each of the five dodecane aggregates simulated, including the aggregation number (n), the aggregate volume (V), the aggregate surface area (A), the aggregate core-minor radius (l_c), the periodic length (PL) of each cylinder and slab, and the degree of curvature (C, as defined in eq 14).

TABLE 3: Overview of Hexadecane MD Simulations[a]

aggregate type	n	V [nm^3]	A [nm^2]	l_c [nm]	PL [nm]	C [nm^{-1}]
small sphere	13	6.31	16.51	1.15		1.74
large sphere	25	12.13	25.53	1.43		1.40
small cylinder	27	13.1	25.19	1.04	3.85	0.96
large cylinder	27	13.1	21.67	1.21	2.85	0.83
slab	96	46.58	28.88	0.81	3.80	0.0

[a] Geometric characteristics of each of the five hexadecane aggregates simulated, including the aggregation number (n), the aggregate volume (V), the aggregate surface area (A), the aggregate core-minor radius (l_c), the periodic length (PL) of each cylinder and slab, and the degree of curvature (C, as defined in eq 14).

$V = 54.3n_{CH_3} + 26.9n_{CH_2}$ (Å^3) where n_{CH_2} is the number of CH_2 groups and n_{CH_3} is the number of CH_3 groups in the aggregate.[6] The surface area, A, of each aggregate was computed geometrically on the basis of the volume of the aggregate and the assumption of a perfectly smooth aggregate surface. The core-minor radius, or planar half-width, l_c, of each aggregate was estimated geometrically using the same assumption. The periodic length, PL, applies only in the case of the cylindrical and the planar aggregates and refers to the periodic simulation cell length of the simulated infinite cylinders, as well as to the periodic simulation cell width and length of the simulated infinite planar layers. Each of the five different geometries simulated for each oil molecule has a different curvature, C, which we have defined using the convention used in the Gibbs−Tolman−Koenig−Buff equation:[37−40]

$$C = \frac{(S-1)}{l_c} \tag{14}$$

where S is a shape factor that is equal to 3 for spheres, 2 for cylinders, and 1 for planar interfaces. As discussed in Section 3.6, although the actual dimensions of the simulated aggregates were chosen arbitrarily, they cover a range of aggregate shapes and sizes that are frequently encountered in modeling micellar systems. Representative spherical, cylindrical, and planar aggregate geometries have been simulated, with values of l_c ranging from 0.81 to 1.17 nm and curvature values ranging from 0 to 1.74 nm^{-1}.

Spherical Oil Aggregates. Each large spherical oil aggregate was prepared by first allowing oil molecules distributed randomly in a simulation cell to self-assemble into a spherical aggregate in vacuum (which is much less computationally expensive to simulate than their self-assembly in water). After self-assembly, which was driven by van der Waals attractions between the oil molecules, sufficient water molecules were added to the simulation cell to ensure that each oil aggregate was at least 2 nm away from its periodic image. An extended *NPT* simulation was then performed during which the *x*-, *y*-, and *z*-dimensions of the simulation cell were allowed to change subject to an applied pressure of 1 bar. During this equilibration period, which was 5 ns for octane and 10 ns for dodecane and hexadecane, both the system potential energy and the solvent accessible surface area (SASA) of each oil aggregate reached a constant value. We consider SASA to be the most important metric to measure equilibration, because this property is directly proportional to the degree of hydration of the oil aggregate, and obtaining accurate hydration information is the key objective of our computer simulations. After equilibration, a 5 ns data-gathering simulation was conducted.

Each small spherical oil aggregate was prepared by first starting from the post-equilibration conformation of a large spherical oil aggregate and then removing half of the oil molecules. An *NPT* equilibration run was subsequently done under these new conditions for 5 ns until both the system potential energy and the SASA of the oil aggregate became constant. During equilibration, the simulation cell dimensions quickly decreased to compensate for the volume of oil that was removed, forming a new simulation cell of approximately the same density as the initial one. Each new oil aggregate constructed in this manner had a radius that was approximately 20% smaller than the original radius. After equilibration, each oil aggregate was simulated for an additional 5 ns under *NPT* conditions during which data was gathered.

Cylindrical Oil Aggregates. Each cylindrical oil aggregate was prepared by first allowing oil molecules distributed randomly in a simulation cell to self-assemble into a cylindrical aggregate in vacuum. After self-assembly, the aggregate was in contact with the periodic boundaries on two sides of the simulation cell (which we define as the two sides perpendicular to the *z*-axis). Next, sufficient water molecules were added to ensure that the cylindrical oil aggregate was always at least 2 nm away from its periodic image in the *x*- and *y*-directions. An extended constant pressure simulation was then performed in which only the *x*- and *y*-dimensions of the simulation cell were allowed to change subject to an applied pressure of 1 bar and during which both the system potential energy and SASA became constant. This equilibration continued for 5 ns for octane and for 10 ns for dodecane and hexadecane. Equilibration was followed by a 5 ns data-gathering simulation.

Initially, the three types of oil molecules were simulated in a simulation cell whose *z*-dimension was 3.85 nm in length. We will refer to this distance as the cylinder "length," although one should keep in mind that, because of the use of periodic boundary conditions, the cylinder is actually infinitely long. To investigate the effect of curvature on the simulation results and to evaluate the effectiveness of the CS−MT modeling approach for cylinders of different curvature, a new simulation cell was constructed for octane, dodecane, and hexadecane by simulating the original simulation cell at constant pressure but without

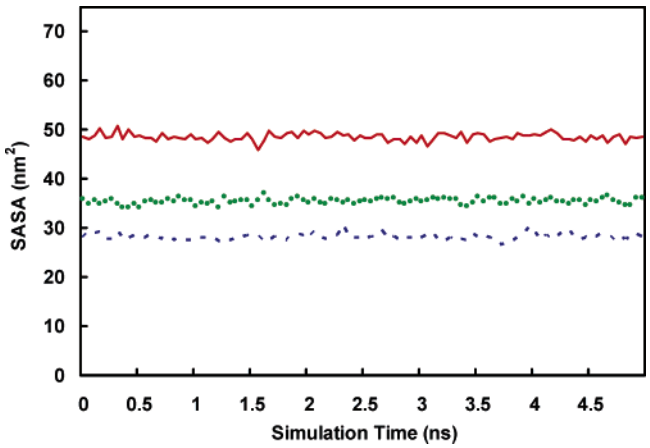


Figure 3. Solvent accessible surface area (SASA) as a function of simulation time for three representative simulated oil aggregates: a small octane sphere (---), a large octane cylinder (●●●), and an octane slab (—). The timescales involved in the SASA fluctuations are much shorter than the total simulation time.

fixing the *z*-dimension. During these constant-pressure simulations, the length of the *z*-dimension of each simulation cell gradually contracted in response to the new boundary conditions. These simulations were continued for 1 to 2 ns until each simulation cell attained a cylinder length of 2.85 nm. At this point, the length of the *z*-dimension of each simulation cell was again fixed. Although each of the three new cylinders formed using this approach had the same volume as the original cylinders, each cylinder core radius was approximately 16% larger than the original cylinder core radius. We will refer to the cylinders that are 3.85 nm in length as the "small" cylinders because they are comparatively thin, and to the cylinders that are 2.85 nm in length as the "large" cylinders because they are comparatively thick. After additional equilibration in the new geometry for 100 ps to ensure that the system potential energy and SASA were stable, each cylindrical oil aggregate was simulated for an additional 5 ns during which data was gathered.

Planar Oil Aggregates. Each planar oil aggregate was prepared by first placing a pre-equilibrated slab of octane, dodecane, or hexadecane within a simulation cell surrounded on two sides by water molecules. After construction, each simulation cell contained two oil−water interfaces perpendicular to what we define as the *z*-axis. For each oil slab, an *NPT* simulation was performed, in which only the *z*-dimension of the simulation cell was allowed to change. This prevented contraction of the simulation cell in the *x*- and *y*-dimensions to reduce the interfacial free energy. Equilibration was evaluated by monitoring the total potential energy of the system and the SASA of each oil slab and ensuring that both properties had stabilized. For octane, a total of 5 ns of equilibration was found to be sufficient. However, for dodecane and for hexadecane, a total of 10 ns was required. After equilibration, a 5 ns data-gathering simulation was conducted.

Equilibration Results. For the three types of oil molecules considered and each of the five aggregate geometries simulated, we observed that the timescales associated with potential energy and SASA equilibration and fluctuation were much shorter than the total simulation time, although SASA fluctuation was found to occur on a much longer time scale than potential energy fluctuation. The total simulation time for each system was sufficient to ensure adequate sampling of both properties. To demonstrate this, in Figure 3, we plot SASA results from the 5 ns data-gathering simulation run for three representative octane aggregate geometries: the small spherical octane aggregate, the large cylindrical octane aggregate, and the planar octane aggregate. As can be seen, the SASA values are stable with time and exhibit no noticeable upward or downward drift during the data-gathering simulation.

Snapshots of the post-equilibration configurations corresponding to (i) the large and small dodecane spherical aggregates, (ii) the large and small dodecane cylindrical aggregates, and (iii) the planar dodecane aggregate are shown in Figure 4. The snapshots of the octane and hexadecane aggregates appear very similar. Each oil molecule is depicted using the van der Waals radius of each atom. For clarity, the water molecules are not shown.

4.5. Data Analysis Method. *4.5.1. Definition of Hydration.* In Sections 1 and 3, we discussed the importance of quantifying the degree of hydration of solutes in the bulk water and aggregate states. Before implementing the new CS−MT modeling approach described in Section 3, we describe below our specific methodology for determining the extent of hydration from the simulation data. In particular, our definition of hydration is based on the number of contacts with "hydrating" atoms, where a hydrating atom is defined as an atom that (i) is capable of forming hydrogen bonds or (ii) is capable of coordinate (dative covalent) bonding. On the basis of this definition, if a hydrophobic CH_2 group is in contact with any atom in a water molecule, with a positively or negatively charged ion, or with a hydrophilic group in the surfactant head that is capable of hydrogen bonding, then that contact is considered "hydrating". We adopt this definition because the hydrophobic effect arises from changes in the hydrogen-bonding or coordinate-bonding network of the aqueous solution that are induced by the presence of nonpolar, hydrophobic moieties.[43] Contact between a hydrophobic group and water, hydrogen-bonding groups in a surfactant head, or a negatively charged ion in solution may break or perturb the hydrogen-bonding network. Similarly, contact between a hydrophobic group and a positively charged ion in aqueous solution may disrupt coordinate bonds between water and the ion. In the CS−MT model, we approximate all hydrating contacts as having the same free energy. The implications of this approximation for modeling nonionic surfactant micellization will be discussed in greater detail here and in subsequent articles on micellization and micellar solubilization.

4.5.2. Analysis of the Bulk Water and the Aggregate Simulation Results. To quantify the degree of hydration of each atom (or group of atoms) in the solute molecule during a bulk water simulation, the number of contacts with hydrogen-bonding or with coordinate-bonding atoms experienced by different atoms in the solute must be counted during the course of a simulation run. For the oil molecules in water considered here, the only contacts that need to be counted as contributing to hydration are contacts with the oxygen and the hydrogen atoms in water. However, for an ionic surfactant in aqueous solution, contacts with water atoms, ions, and hydrogen-bonding groups in the surfactant head should each be counted as contributing to hydration. In analyzing our simulation data, a contact was defined as two atoms being separated by less than a set distance (the "cutoff" distance) at any time during the simulation. It is important to note that the average number of contacts counted using this method of analysis is directly proportional to the average number of hydrating atoms located within the specified cutoff distance.

Quantifying the degree of hydration of the solute molecules during the aggregate simulations was done in the same manner (and using the same cutoff distance used to identify contacts)

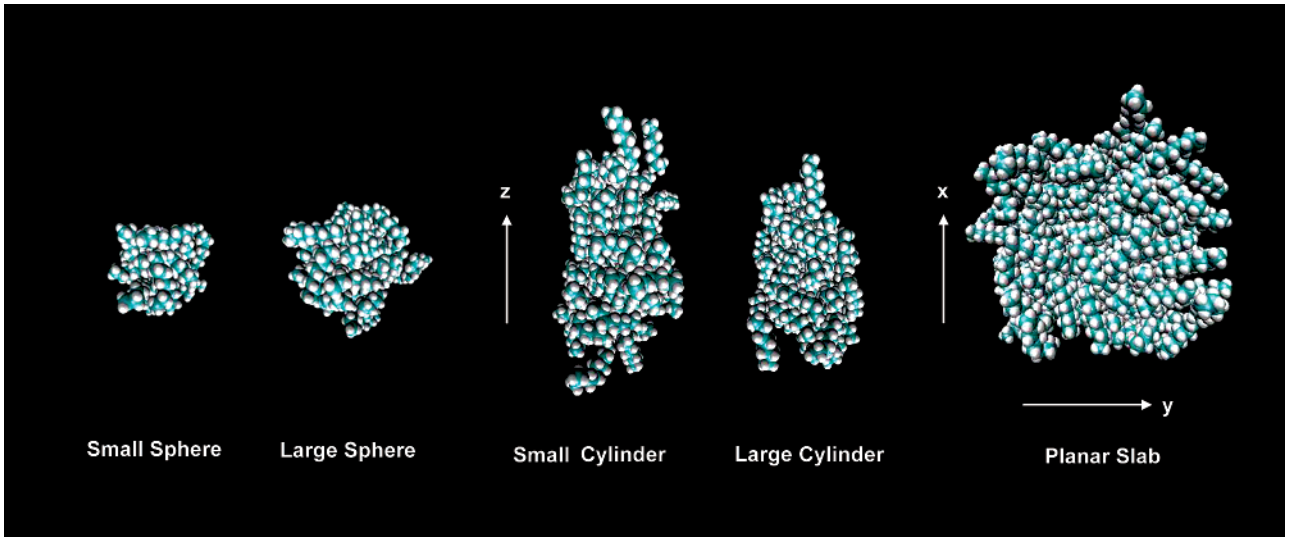


Figure 4. Snapshots of the post-equilibration structures of (i) two dodecane spherical aggregates (of aggregation numbers 17 and 33), (ii) two dodecane cylindrical aggregates (of periodic lengths 2.85 and 3.85 nm), and (iii) a dodecane slab. The water molecules are not shown for clarity.

as was done in analyzing the results of the bulk water simulations. After counting the hydrating contacts in the bulk water and in the aggregate states, we analyzed the hydration data using the metric introduced to quantify hydration in Section 3. Recall that the relative degree of hydration, f, is defined as the extent of hydration in the aggregate environment divided by the extent of hydration in the bulk water environment (see eq 10). On the basis of this definition, f values for each solute molecule were computed from the simulation data as follows:

$$f = \frac{\text{number of hydrating contacts in the aggregate}}{\text{number of hydrating contacts in the bulk water}} \quad (15)$$

As expected intuitively, we found that f is significantly less than unity for a hydrophobic solute, because fewer contacts with hydrating atoms are experienced in the aggregate environment than in the bulk water environment.

4.5.3. Selecting the Cutoff Distance. In selecting the cutoff distance used to define contacts between atoms, we were guided by the realization that, in implementing the CS−MT model, we are interested only in quantifying the *local* environment of each hydrophobic atom or group of atoms. Accordingly, a cutoff distance should be selected such that only nearest-neighbor atoms contribute contacts to a hydrophobic group. However, to ensure that good contact statistics are obtained, the cutoff distance selected should be at least as large as the sum of the van der Waals radii of two hydrogen atoms (one that is bonded to a CH_2 or to a CH_3 group and the other that is bonded to a water molecule) or 0.24 nm.

To determine the sensitivity of the CS−MT modeling results to the value of the cutoff distance selected, several different cutoff values were tested, including 0.25, 0.3, 0.4, and 0.5 nm. Note that when computing f using eq 15, the same cutoff value was used to count contacts in both the bulk water and the aggregate environments. The CS−MT modeling results were found to be weakly dependent on the value of the cutoff distance when modeling planar and curved oil aggregates. We found that by choosing the smallest value of the cutoff distance that yields good statistics, only nearest-neighbor contacts with hydrating atoms were included, and the dependence of f on curvature was minimized. As discussed in Section 3, the effect of curvature on g_{form} is accounted for theoretically in the CS−MT model and need not be included by using a large cutoff distance that introduces curvature dependence into the calculated f values. We also computed radial distribution functions between the hydrophobic CH_2 and CH_3 groups and water using the bulk water simulation data to aid in determining an appropriate cutoff distance. On the basis of the results of our sensitivity and radial distribution function tests (results not reported), we selected a cutoff distance of 0.3 nm as being most appropriate when implementing the CS−MT model. All the CS−MT modeling results reported in this article were generated using this cutoff distance.

4.5.4. Error Analysis. An estimate of the standard error in f for each group of atoms in the solute molecules was obtained through the use of block averaging.[60−62] In block averaging, the standard error is computed from the variance between averages of blocks of data, and the block size is increased until the standard error estimate becomes constant. To assist in identifying this asymptotic value for the simulation data reported here, a two-exponential function was fit to the block average curve.[60−62] Block averaging is useful to analyze correlated data, such as the results obtained from a MD simulation. Data-gathering simulation runs for solute molecules in the bulk water and in the aggregate states were conducted for sufficient time to ensure that the uncertainty in each calculated value of f was sufficiently small—typically less than 5%.

The block averaging approach described above provides an accurate estimate of the standard error of the results of a single simulation. However, typically, it is also desirable to run multiple independent simulations to estimate the run-to-run variance. If the run-to-run variance is much larger than the variance estimated from a single simulation, it indicates that insufficient sampling has been done.[51,54,56] This problem has been commented upon in the context of free-energy calculations using computer simulation.[53,63] Although the block averaged results for individual runs presented here indicate a high degree of statistical certainty, we conducted additional independent bulk water and aggregate simulations to determine the run-to-run variance. The run-to-run variance was found to be comparable in size to the block average estimates of the standard error for each solute (results not reported). Accordingly, and because of

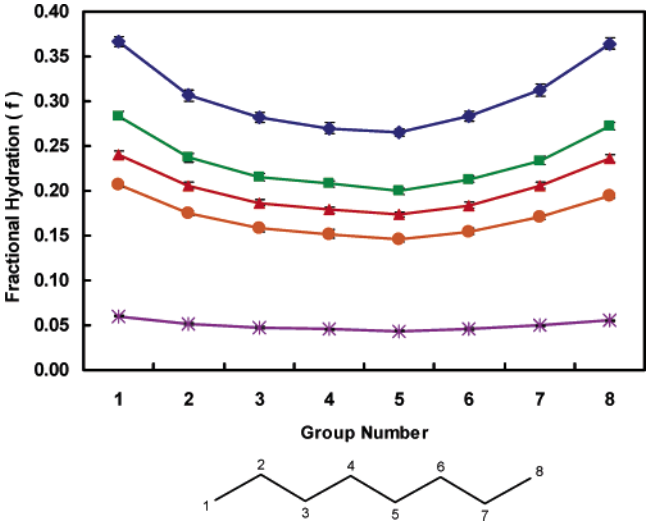


Figure 5. The average fractional degree of hydration (f), as defined in eq 15, corresponding to each of the groups in octane for each of the five simulated aggregate geometries. Results are presented for (i) a small spherical aggregate (−◆−), (ii) a large spherical aggregate (−■−), (iii) a small cylindrical aggregate (−▲−), (iv) a large cylindrical aggregate (−●−), and (v) a planar slab (−*−). Group numbers are identified in the molecular structure shown below the fractional hydration plot. The error bars shown correspond to the standard error of the mean. The various lines are shown as a guide for the eye.

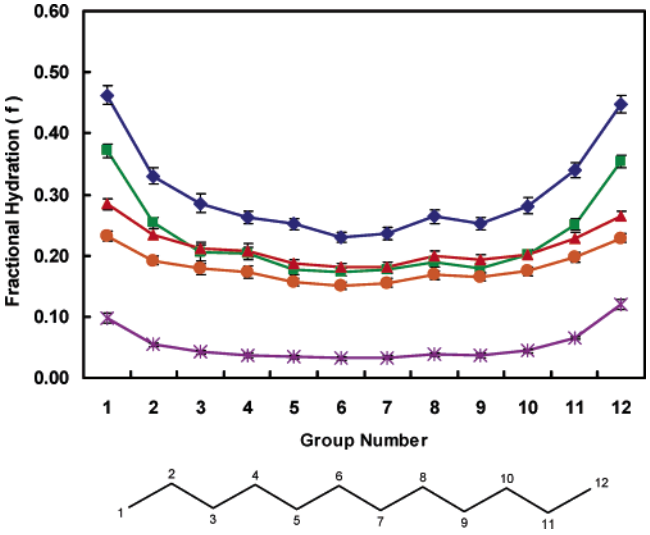


Figure 6. The average fractional degree of hydration (f), as defined in eq 15, corresponding to each of the groups in dodecane for each of the five simulated aggregate geometries. Results are presented for (i) a small spherical aggregate (−◆−), (ii) a large spherical aggregate (−■−), (iii) a small cylindrical aggregate (−▲−), (iv) a large cylindrical aggregate (−●−), and (v) a planar slab (−*−). Group numbers are identified in the molecular structure shown below the fractional hydration plot. The error bars shown correspond to the standard error of the mean. The various lines are shown as a guide for the eye.

the high computational cost associated with conducting such simulations, independent simulations were not considered to be necessary for the surfactant systems that will be simulated in article 2 of this series.

5. Simulation Results and Discussion

5.1. Fractional Hydration Results. Simulations of a single molecule of octane, dodecane, and hexadecane in bulk water were conducted to determine the average number of contacts experienced by each of the CH_2 and the CH_3 groups in the oil molecules with water in the bulk aqueous state. Subsequently, aggregate simulations were conducted for each oil type and for the five different aggregate geometries discussed in Section 4.4.2.

The resulting average fractional degree of hydration, f, is plotted as a function of group number for octane in Figure 5, for dodecane in Figure 6, and for hexadecane in Figure 7. The error bars shown represent the standard error of the mean for each value of f and are typically of the size of the various symbols shown or smaller. As discussed in Section 4.4.2, to the extent possible (given the requirement of an integer number of molecules), aggregation numbers were selected such that corresponding aggregates of octane, dodecane, and hexadecane have the same dimensions. Therefore, the simulation results presented here for each oil molecule can be compared directly to evaluate the effect of the hydrocarbon chain length on f and SASA.

As shown in Figures 5−7, each of the three oil molecules considered has a symmetric degree of hydration profile in which each structurally equivalent CH_2 and CH_3 group has the same degree of hydration within the error of the simulation results. The CH_3 groups on both ends of each oil molecule have the highest f values, and the CH_2 groups near the middle of each oil molecule have the lowest f values. The average value of f for each oil aggregate is related to the exposed surface area per molecule, a. For a perfectly smooth oil aggregate, $a = Sv_t/l_c$,

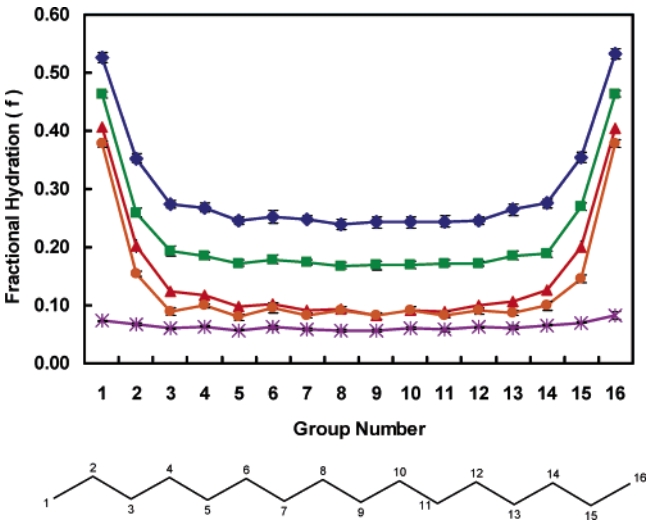


Figure 7. The average fractional degree of hydration (f), as defined in eq 15, corresponding to each of the groups in hexadecane for each of the five simulated aggregate geometries. Results are presented for (i) a small spherical aggregate (−◆−), (ii) a large spherical aggregate (−■−), (iii) a small cylindrical aggregate (−▲−), (iv) a large cylindrical aggregate (−●−), and (v) a planar slab (−*−). Group numbers are identified in the molecular structure shown below the fractional hydration plot. The error bars shown correspond to the standard error of the mean. The various lines are shown as a guide for the eye.

where v_t is the volume of the solute tail.[7] Although the various oil aggregates simulated here are not perfectly smooth, we have found that the inverse relationship between a and l_c is still valid. The relatively low values of f for the planar oil aggregates (see the −*− results in Figures 5−7), for example, can be understood by noting that this aggregate has a value of a that is approximately 4.3 times larger than that of the small spherical aggregate (−◆−), 3.5 times larger than that of the large spherical aggregate (−■−), 3.2 times larger than that of the small

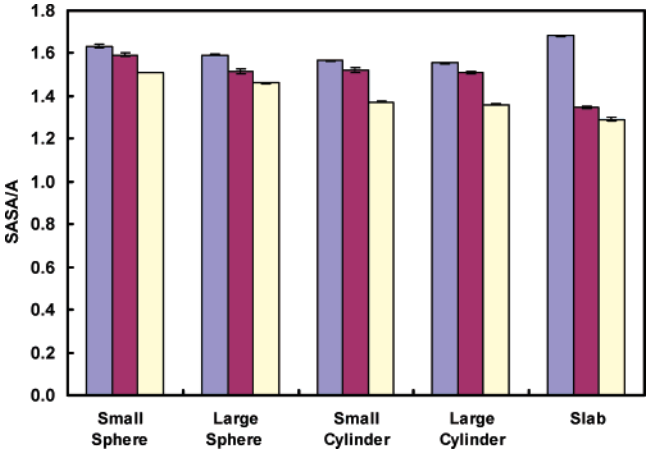


Figure 8. The ratio of the solvent accessible surface area (SASA) to the geometrically defined surface area (*A*, see text) for each of the octane (blue), dodecane (maroon), and hexadecane (yellow) oil aggregates simulated, grouped according to geometry. The error bars correspond to the standard error of the mean.

cylindrical aggregate (−▲−), and 2.7 times larger than that of the large cylindrical aggregate (−●−).

It is interesting to note that, although the average value of *f* for oil molecules in aggregates of the same shape and size are approximately the same for each of the three oil molecules, the magnitude of the difference between the highest and the lowest *f* value in each molecule is different for the three oil molecules. For example, the average magnitude of this difference for all five hexadecane geometries is 73% larger than the average magnitude of this difference for all five octane geometries, it is 53% larger for all five dodecane geometries than for all five octane geometries, and it is 43% larger for all five hexadecane geometries than for all five dodecane geometries. Visual inspection of the trajectories for each aggregate simulation revealed that these differences are due to a packing effect. More specifically, both dodecane and hexadecane exhibited relatively high degrees of hydrocarbon chain alignment after equilibration. This alignment resulted in a disproportionate number of CH_3 groups being adjacent to the aggregate core−water interface rather than being buried in the aggregate interior, which in turn imparts relatively high *f* values to the terminal CH_3 groups.

5.2. Solvent Accessible Surface Area (SASA) Results. The values of *f* computed for octane, dodecane, and hexadecane are intimately related to the average value of SASA associated with each oil aggregate interface. We have computed average values of SASA for each of the 15 oil aggregate geometries considered on the basis of the simulation results. To calculate SASA, we used the double cubic lattice method as implemented in GROMACS. The solvent accessible surface was traced out by a probe sphere of radius 0.2 nm that was rolled around each molecule within the aggregate to identify the solvent accessible regions.[55] This probe sphere radius was selected on the basis of the size of a water molecule (for which a probe radius of 0.14 nm is frequently used) and the requirement of preventing the probe from identifying any of the aggregate core region as being solvent accessible. The values of the time-averaged SASA divided by *A*, the perfectly smooth surface areas reported in Tables 1−3, are shown in Figure 8. Note that the ratio, SASA/*A*, provides a convenient measure of surface roughness. The error bars shown in Figure 8 represent the standard error of the mean. As can be seen, the value of SASA/*A* for the 5 geometries considered is equal to 1.61 for octane (blue), 1.50 for dodecane (maroon), and 1.40 for hexadecane (yellow). It is interesting to note that, for each geometry, SASA/*A* is lower for hexadecane than for dodecane and for dodecane than for octane. Our results indicate that, as the alkyl chain length decreases, the average roughness of the surface of the oil aggregate increases. In addition, a comparison of the results for the small and large spheres, and the results for the small and large cylinders, reveals that, for each oil type considered, SASA/*A* is higher for the small aggregates and lower for the large aggregates.

As discussed in Section 3.4, to calculate g_{hydr}, we proposed a theoretical approach to estimate Δg_{wc} using eq 13. Recall that Δg_{wc} is equal to $\sigma_{core} - \sigma_{bulk}$. As such, it represents the free-energy difference (on a per unit SASA basis) associated with hydrating contacts in the aggregate state and the bulk water state. The value of σ_{core} is a function of both the alkyl tail length and the hydrophobic core curvature (see Section 3.4). We can estimate $\sigma_{core} = \sigma A_{core}/SASA_{core}$ using a value of σ corresponding to an oil molecule of type *j* (*j* = octane, dodecane, or hexadecane) calculated using eq 7. The ratio, $A_{core}/SASA_{core}$, is simply equal to the inverse of the SASA/*A* ratios reported in Figure 8. We note that, because $SASA_{core}$ is significantly larger than A_{core}, σ_{core} is significantly smaller than σ. We have performed a linear regression on the data to describe $SASA_{core}/A_{core}$ (or the "roughness" of the oil aggregate−water interface) as a function of linear alkyl chain length (n_t) and oil aggregate curvature (*C*, as defined in eq 14). The resulting expression, which has been fit with an R^2 value of 0.76, is given by

$$SASA_{core}/A_{core} = 1.740 - 0.026n_t + 0.078C \quad (16)$$

Although σ_{core} is different for each oil aggregate, σ_{bulk} is approximately constant (see Section 3.4). Using a 0.2 nm probe for water (to be consistent with the SASA estimates for the oil aggregates), we computed $SASA_i$ values for octane, dodecane, and hexadecane and then used these values to compute $\sigma_{bulk} = -g_{tr_i}/SASA_i$ using known values of g_{tr_i} (see Section 3.3). Our computed values of σ_{bulk} for octane, dodecane, and hexadecane are within 6% of each other, demonstrating that modeling the hydration free energy as being proportional to SASA for solutes of the same size as that of typical surfactant tails is a reasonable approximation. In addition, the average estimate of σ_{bulk} that we obtained for octane, dodecane, and hexadecane is 26.84 cal/mol/Å^2 (using a 0.2 nm probe for water to be consistent with the SASA estimates presented for the oil aggregate). As such, our result for σ_{bulk} is similar to σ_{bulk} estimates given by Tanford that are between 20 and 25 cal/mol/Å^2.[44]

We have predicted Δg_{wc} using (i) the actual $SASA_{core}/A_{core}$ values obtained from our computer simulation results and (ii) the correlation for $SASA_{core}/A_{core}$ given in eq 16 for each of the 15 oil aggregates considered. Both predicted values of Δg_{wc} are plotted versus the oil aggregate curvature in Figure 9, where the curvature is defined in eq 14. As can be seen, the agreement between the two theoretical estimates of Δg_{wc} is reasonable. In general, our results indicate that Δg_{wc} decreases with increasing curvature. In addition, the change in Δg_{wc} with respect to curvature is smallest for octane (see the ♦ results) and largest for hexadecane (▲).

6. Molecular-Thermodynamic Modeling Based on Computer Simulation Inputs

We next use the CS−MT model to calculate g_{form} for each of the 15 oil aggregates discussed in Section 5. As discussed in Section 3.6, we have chosen oil aggregates as a starting point to evaluate the validity and accuracy of the CS−MT model because these structures do not require dealing with the computational challenges posed by the presence of the surfactant

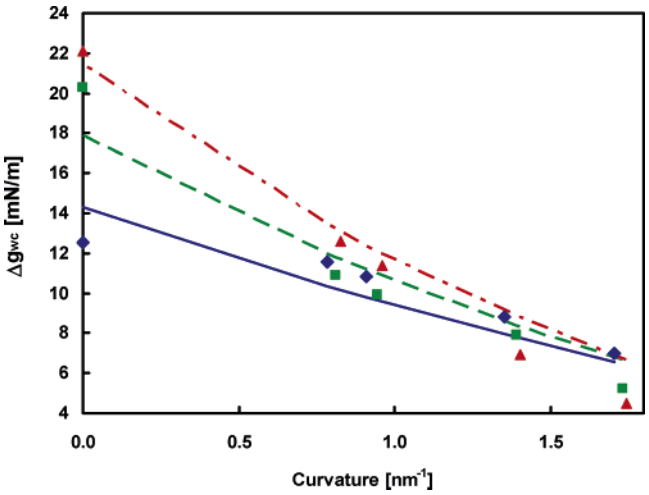


Figure 9. Theoretical predictions of the hydration free-energy difference (Δg_{wc} [mN/m]) associated with hydrating contacts in the aggregate state and bulk water state are shown as a function of curvature (nm^{-1}), where the curvature, C, is defined using eq 14. Theoretical predictions based on values of $SASA_{core}/A_{core}$ computed from the simulation results are reported for octane (◆), dodecane (■), and hexadecane (▲). Theoretical predictions based on values of $SASA_{core}/A_{core}$ obtained using the correlation given in eq 16 are also reported for octane (—), dodecane (– –), and hexadecane (– ·).

heads. Because experimental data is not available for the oil aggregates considered (because they are not thermodynamically stable except at infinite dilution), we compare our CS−MT model predictions with the predictions of the traditional MT model.

As discussed in Section 3.6, for each of the simulated oil aggregates, there are only two nonzero contributions to the free energy of aggregate formation, g_{form}. These contributions are g_{tr} and g_{int} in the traditional MT modeling approach and g_{dehydr} and g_{hydr} in the CS−MT modeling approach.

CS−MT model predictions for g_{form} were made using eq 9 (where g_{pack}, g_{st}, g_{elec}, and g_{ent} are all equal to 0) and a 0.3 nm cutoff for the identification of the water contacts. The free energy of dehydration, g_{dehydr}, and the free energy of hydration, g_{hydr}, were computed using eqs 11 and 12, respectively. In computing g_{hydr}, we used the theoretical model for Δg_{wc} given in eq 13 and the correlation for $SASA_{core}/A_{core}$ given in eq 16. Traditional MT model predictions for g_{form} were made using eq 4 (where g_{pack}, g_{st}, g_{elec}, and g_{ent} are all equal to 0) and by combining eqs 5−7 to obtain the following expression:

$$g_{form} = \sum_{i=1}^{n_{core}} g_{tr_i} + a\left[\frac{\sigma_{0,j}}{\left(1 + \frac{(S-1)\delta}{l_c}\right)}\right] \qquad (17)$$

In Table 4, we report predictions of the CS−MT model and the traditional MT model for g_{form}, g_{dehydr}, and g_{hydr} corresponding to each of the five octane, dodecane, and hexadecane aggregate geometries simulated and modeled in this article. Each of the reported errors for the CS−MT model predictions corresponds to the standard error of the mean and were computed through block averaging of the simulation results in the manner described in Section 4.5.4. Each reported Δ value corresponds to the percent difference between the CS−MT model and the MT model predictions for g_{form}. The estimated uncertainty in the CS−MT model predictions is comparable in magnitude to the difference between the predictions of the CS−MT model and those of the traditional MT model.

The average absolute discrepancy between the predictions of the CS−MT model and those of the traditional MT model for g_{form}, or the average of the absolute value of the 15 Δ values reported in Table 4, is only 1.04%. The best agreement between the predictions of the CS−MT model and those of the traditional MT model is for the small cylinders, with an average Δ value of only −0.05%. This is followed by the large sphere results (average Δ value = −0.56%), the large cylinder results (average Δ value = 0.74%), the slab results (average Δ value = 1.03%), and finally the small sphere results (average Δ value = −1.53%). On average, the differences are negative for the spheres and positive for the cylinders and the slabs, but in all cases, the errors are small. It is interesting to note that although, in all cases, the magnitude of g_{hydr} is much smaller than that of g_{dehydr}, it must be included in the CS−MT model for g_{form} in order to yield such a high level of agreement between the CS−MT model and the traditional MT model. Given the fact that the value of Δg_{wc} used to compute g_{hydr} in the CS−MT modeling approach was predicted theoretically using eqs 13 and 16, rather than through any type of fitting procedure, the very high level of agreement between the predictions of the CS−MT model and the traditional MT model further supports the validity of the proposed CS−MT modeling approach. In addition, it supports the validity of the procedure that we adopted to define and count hydrating contacts as well as to compute SASA values in implementing the new CS−MT modeling approach.

For the completely hydrophobic solutes (oils) modeled in this article, the equivalence of the CS−MT model and the traditional MT model for the hydrophobic effect can be demonstrated mathematically if the following two conditions are met: (i) each f_i value must be defined so that it only conveys information about the *local* environment of group i (i.e., the environment immediately surrounding the solvent accessible surface of group i), and (ii) the approach used to determine $SASA_i$ and $SASA_{core}$ must yield physically realistic results, where the same solute probe size must be used to quantify both values. A demonstration of the mathematical equivalence of the CS−MT model and the traditional MT model is given in Appendix B, along with a discussion of the criteria that must be met for the two models to be equivalent. The close agreement between the results of the CS−MT model and the traditional MT model for the 15 oil aggregates considered here demonstrates that criteria (i) and (ii) above are indeed satisfied and that the simulation times used to gather hydration data were sufficient to provide highly accurate values of f_i. It is important to note that, for amphiphilic solutes such as surfactants (which possess both a head and a tail), the CS−MT model and the traditional MT model are not equivalent even if criteria (i) and (ii) are met. The reasons behind this lack of equivalence are discussed in detail in Appendix B. We believe that the CS−MT model yields more realistic estimates of the hydrophobic driving force for micelle formation in the case of complex amphiphilic solutes. With this expectation in mind, in article 2 of this series, we will implement the CS−MT model in the case of nonionic surfactants and will also compare predictions made by the CS−MT model with those made by the traditional MT model for these surfactants.

Using the data reported in Table 4, we found that it was possible to evaluate the validity of a computational approach that we detail in Appendix A to extend the applicability of the CS−MT model to allow the prediction of g_{form} as a function of aggregate shape and size, rather than only for a specific simulated aggregate geometry. Briefly, in Appendix A, we will show that the CS−MT model can be used to predict the optimal

TABLE 4: CS−MT and Traditional MT Modeling Results[a]

octane					
aggregate type	g_{dehydr} [k_BT]	g_{hydr} [k_BT]	g_{form} [k_BT]		
			CS−MT	MT	Δ (%)
small sphere	−10.86 ± 0.04	1.82 ± 0.01	−9.04 ± 0.04	−9.13	−1.08
large sphere	−12.09 ± 0.03	1.74 ± 0.01	−10.35 ± 0.03	−10.31	0.42
small cylinder	−12.63 ± 0.02	2.03 ± 0.02	−10.60 ± 0.03	−10.59	0.06
large cylinder	−13.16 ± 0.02	1.66 ± 0.01	−11.5 ± 0.02	−11.27	1.97
slab	−15.14 ± 0.01	0.60 ± 0.00	−14.53 ± 0.01	−14.12	2.84
dodecane					
aggregate type	g_{dehydr} [k_BT]	g_{hydr} [k_BT]	g_{form} [k_BT]		
			CS−MT	MT	Δ (%)
small sphere	−14.69 ± 0.09	1.94 ± 0.02	−12.75 ± 0.09	−12.85	−0.74
large sphere	−16.42 ± 0.07	2.23 ± 0.03	−14.19 ± 0.07	−14.24	−0.37
small cylinder	−17.01 ± 0.06	2.52 ± 0.03	−14.49 ± 0.07	−14.41	0.56
large cylinder	−17.80 ± 0.05	2.33 ± 0.03	−15.47 ± 0.06	−15.32	0.96
slab	−20.59 ± 0.04	1.41 ± 0.04	−19.18 ± 0.06	−19.09	0.46
hexadecane					
aggregate type	g_{dehydr} [k_BT]	g_{hydr} [k_BT]	g_{form} [k_BT]		
			CS−MT	MT	Δ (%)
small sphere	−18.62 ± 0.06	2.14 ± 0.01	−16.48 ± 0.07	−16.94	−2.78
large sphere	−20.73 ± 0.05	2.56 ± 0.02	−18.17 ± 0.05	−18.48	−1.72
small cylinder	−22.68 ± 0.04	3.28 ± 0.02	−18.31 ± 0.05	−18.45	−0.78
large cylinder	−23.24 ± 0.04	2.92 ± 0.02	−19.39 ± 0.05	−19.53	−0.73
slab	−26.12 ± 0.02	2.10 ± 0.03	−24.01 ± 0.03	−24.06	−0.19

[a] Computer simulation−molecular-thermodynamic (CS−MT) and traditional molecular-thermodynamic (MT) modeling results for g_{dehydr}, g_{hydr}, and g_{form} corresponding to each of the five octane, dodecane, and hexadecane aggregate geometries simulated and modeled in this article. The uncertainties reported for the CS−MT model results correspond to the standard error of the mean. Each Δ value presented in the table represents the percent difference between the CS−MT model and the traditional MT model predictions for g_{form}.

shape and size of solute aggregates based on fractional hydration information obtained from simulation of a *single* aggregate.

7. Conclusions

We have developed a novel computer simulation−molecular-thermodynamic (CS−MT) modeling approach that allows one to eliminate many of the simplifying assumptions that were needed to quantify the magnitude of the hydrophobic effect in the traditional MT modeling approach. This was accomplished by making use of detailed hydration data obtained through atomistic computer simulations. The detailed hydration information was used in a new theoretical framework to quantify the hydrophobic free-energy contributions associated with the self-assembly of hydrophobic solutes. In this model, g_{form} is computed using the expression $g_{form} = g_{dehydr} + g_{hydr} + g_{pack} + g_{st} + g_{elec} + g_{ent}$, where g_{dehydr} is the dehydration free-energy contribution and g_{hydr} is the hydration free-energy contribution. These two free-energy contributions replace the transfer, g_{tr}, and the interfacial, g_{int}, free-energy contributions, which quantify the hydrophobic effect in the traditional MT modeling approach. The remaining free-energy contributions (g_{pack}, g_{st}, g_{elec}, and g_{ent}) are calculated in the context of the new CS−MT modeling approach in the same manner that they are calculated in the traditional MT modeling approach. However, it is important to stress that computer simulation information obtained on the aggregate structure may be used to improve the estimation of g_{pack} and g_{st}. Work along these lines is in progress.

The free-energy contribution associated with dehydration, g_{dehydr}, is modeled using the concept of a transfer free energy for each individual hydrophobic group in a solute. These transfer free energies can be obtained from experimental solubility data or estimated theoretically using group-contribution methods. The free-energy contribution associated with hydration, g_{hydr}, accounts for the change in free energy associated with water contacts in the aggregate and in the bulk water states. In this article, we have proposed a specific method to theoretically calculate this free-energy difference in the case of oil molecules. However, the CS−MT model was formulated in a general way and may be used to model single nonionic, zwitterionic, and ionic surfactant micellization, the micellization of mixtures of nonionic and ionic surfactants, and micellar solubilization in addition to the self-assembly of oil aggregates. Although the theoretical model for Δg_{wc} presented in this article was formulated only on the basis of data obtained from oil aggregate simulations, we anticipate that it may be successfully used to model the self-assembly of a variety of solutes and solute mixtures. The free energy per unit SASA in the bulk water reference state, σ_{bulk}, for a hydrophobic group in a surfactant molecule should be very similar to σ_{bulk} for a hydrophobic group in an oil molecule. In addition, the success of the traditional MT modeling approach in modeling the aggregate core−water interface using an oil−water interfacial tension (see Section 2.3.2) does indicate that approximating the free energy per unit SASA in the aggregate reference state, σ_{core}, as being equal to σ_{core} for an oil aggregate should yield reasonable results.

In this article, we have demonstrated the validity and the accuracy of the new CS−MT model by using it to model spherical, cylindrical, and planar oil aggregates, each containing three different types of oil molecules and having different degrees of curvature. Excellent agreement between the predictions of the CS−MT model and those of the traditional MT model for g_{form} was obtained for each of the 15 oil aggregates considered, with an average absolute error of only 1.04% between the two theoretical approaches. Our results also demonstrate that the CS−MT model can be used to predict g_{form} for aggregates of arbitrary shapes and sizes by using hydration information obtained using only two independent molecular dynamics simulations.

Because the hydrophobic effect is the primary driving force for micelle formation in aqueous solution and also the largest free-energy contribution to g_{form}, the modeling approach presented in this article provides an important new approach to more accurately model the self-assembly of hydrophobic and amphiphilic solutes in aqueous solution. By quantifying the actual hydration changes that occur upon self-assembly for various moieties within a solute, and by subsequently using this information in a new theoretical model to quantify the hydrophobic effect, the CS−MT modeling approach has the potential to improve our ability to model the micellization and the micellar solubilization behavior of complex surfactants and solubilizates in aqueous solution. In article 2 of this series,[26] we will use the CS−MT modeling approach presented here to model the micellization behavior of nonionic surfactants in aqueous solution, and in article 3 of this series,[27,28] we will report modeling results for the aqueous micellization of ionic and zwitterionic surfactants in aqueous solution.

Acknowledgment. This research was supported in part by funding provided by DuPont through the DuPont−MIT Alliance. B.C.S. and A.G. are also grateful to Dr. Isaac Reif for critically reading the manuscript and for stimulating discussions about the CS−MT model.

Appendix A: Extension of the CS−MT Model to Predict Aggregate Shape and Size

As discussed in Section 3.5, a key capability of traditional MT modeling is that it enables prediction of g_{form} as a function of aggregate shape and size. From this functional dependence, the optimal aggregate shape and size can be predicted. In this appendix, we outline a computational approach to extend the CS−MT modeling approach to enable prediction of g_{form} as a function of aggregate shape and size. It is important to note that the MD simulations conducted as part of the CS−MT modeling approach do not allow direct prediction of the optimal aggregate shape and size because the simulation timescales are too short to permit the exchange of solute monomers between the aggregate phase and the aqueous phase.

With the exception of the transfer free-energy contribution, g_{tr}, the other five free-energy contributions to g_{form} in eq 4 are all functions of the micelle shape and size.[14] Note that g_{tr} is independent of the structural characteristics of the aggregate because it corresponds to the free-energy change associated with transferring a solute tail from infinite dilution in aqueous solution to an infinite bulk phase composed of solute tails. In order for the CS−MT model to be able to predict micelle shape and size, we will exploit this property of g_{tr}. Specifically, we combine the traditional MT model for the hydrophobic effect ($g_{tr} + g_{int}$) with the CS−MT model for the hydrophobic effect ($g_{dehydr} + g_{hydr}$) to obtain a CS−MT model prediction for the transfer free energy of a solute. Specifically,

$$g_{tr,CS-MT} = g_{dehydr} + g_{hydr} - \hat{g}_{int} \quad \text{(A1)}$$

where $g_{tr,CS-MT}$ is the transfer free energy computed using the CS−MT modeling approach and $\hat{g}_{int}$ is the MT prediction of the interfacial free energy of the simulated micellar aggregate. The free energy of aggregate formation, g_{form}, for micelles of a different shape and size than those for which the computer simulation data was collected can then be calculated using eq 4 in the context of the traditional MT modeling approach:

$$g_{form} = g_{tr,CS-MT} + g_{int} + g_{pack} + g_{st} + g_{elec} + g_{ent} \quad \text{(A2)}$$

TABLE A1: Predicted Values of $g_{tr,CS-MT}$ for Octane, Dodecane, and Hexadecane[a]

	octane	dodecane	hexadecane
small sphere	-15.90 ± 0.04	-21.88 ± 0.09	-27.49 ± 0.07
large sphere	-16.04 ± 0.03	-21.92 ± 0.07	-27.64 ± 0.05
small cylinder	-16.01 ± 0.03	-22.06 ± 0.07	-27.81 ± 0.05
large cylinder	-16.23 ± 0.02	-22.13 ± 0.06	-27.81 ± 0.05
slab	-16.41 ± 0.01	-22.06 ± 0.06	-27.91 ± 0.03
mean ± SD	-16.12 ± 0.20	-22.01 ± 0.10	-27.73 ± 0.16

[a] The results (in k_BT) were obtained using eq A1, and computer simulation data obtained for each of the five aggregate geometries considered. The uncertainty reported for each value of $g_{tr,CS-MT}$ corresponds to the standard error of the mean, as obtained through block averaging (see section on error analysis). The mean values reported for octane, dodecane, and hexadecane in the last row of the table are the average estimates for $g_{tr,CS-MT}$ obtained from the results for each oil type, and the standard deviation (SD) values are the standard deviations in the five $g_{tr,CS-MT}$ estimates obtained for each oil type.

After computing g_{form} using eq A2, the optimal values of S, l_c, α, and β can be obtained in a computationally efficient manner by minimizing g_{form} with respect to each of these variables (see Section 2.2). Using this strategy, the CS−MT modeling approach can be used to predict g_{form} for micelles of any shape and size using only two independent computer simulations: one simulation of the solute in the bulk water environment and a second simulation of the solute in an aggregate environment of arbitrary shape and size.

We would like to stress that, for the proposed computational approach to be implemented successfully, during the aggregate simulation, the aggregate must remain stable and not break apart into monomers or into several smaller aggregates. This stability, of course, is not guaranteed when a micelle is preformed and simulated at a nonoptimal shape and size. Fortunately, on the basis of our experience, surfactant micelles in a somewhat nonoptimal geometry (for example, a spherical instead of a cylindrical geometry, or at a nonoptimal aggregation number) do remain stable during the 10 to 25 ns simulations conducted in the context of the CS−MT modeling approach. As will be discussed in article 2 of this series, each nonionic surfactant micelle that we simulated remained stable on these timescales, even when preformed at a nonoptimal shape and size. In addition, we note that all the 15 oil aggregates that we simulated in this article remained stable during the equilibration and the data-gathering simulation runs.

Using the CS−MT model predictions for g_{form} presented in Table 4, we computed values of $g_{tr,CS-MT}$ for octane, dodecane, and hexadecane by evaluating eq A1 using hydration information obtained from each of the five simulated aggregate geometries. Results for $g_{tr,CS-MT}$ calculated in this manner are reported in Table A1. For the computational approach outlined here to be successful, similar estimates of $g_{tr,CS-MT}$ should be obtained for octane, dodecane, and hexadecane regardless of which aggregate simulation was used to obtain the hydration data.

As Table A1 shows, the predicted values of $g_{tr,CS-MT}$ for octane, dodecane, and hexadecane are quite consistent. The mean and standard deviation of the $g_{tr,CS-MT}$ values predicted for each oil type are reported in the last row of Table A1. The standard deviations reported for each oil type are quite small (between 0.10 and $0.20k_BT$), a result that supports the validity of the computational approach presented here to predict g_{form} for aggregates of arbitrary shapes and sizes.

To illustrate the manner in which eq A2 may be used to reduce the number of simulations required in the CS−MT modeling approach, we have used it to predict g_{form} for each of

TABLE A2: Predicted Values of g_{form} for the Five Simulated Geometries of Dodecane[a]

	g_{form} [k_BT]	$g_{form,max}$ [k_BT]	$g_{form,min}$ [k_BT]	$\|\|\Delta$ (%) $\|\|$ [k_BT]
small sphere	-12.75 ± 0.09	-12.75 ± 0.09	-12.93 ± 0.09	1.43
large sphere	-14.19 ± 0.07	-14.14 ± 0.07	-14.33 ± 0.07	1.29
small cylinder	-14.49 ± 0.07	-14.32 ± 0.07	-14.50 ± 0.07	1.27
large cylinder	-15.47 ± 0.06	-15.22 ± 0.06	-15.40 ± 0.06	1.19
slab	-19.18 ± 0.06	-19.00 ± 0.06	-19.18 ± 0.06	0.96

[a] CS−MT modeling results obtained using eq 9 are reported as g_{form} (in k_BT). CS−MT modeling results obtained using eq A2 and the maximum value of $g_{tr,CS-MT}$ for dodecane reported in Table 4 are reported as $g_{form,max}$ (in k_BT). CS−MT modeling results obtained using eq A2 and the minimum value of $g_{tr,CS-MT}$ for dodecane reported in Table 4 are reported as $g_{form,min}$ (in k_BT). The uncertainty reported for each value corresponds to the standard error of the mean, as obtained through block averaging (see section on error analysis).

the five dodecane aggregates considered in this article. The results are reported in Table A2. The first column of predictions in Table A2 for the free energy of formation (reported as g_{form}) is identical to what is reported in Table 4 as the CS−MT model predictions of g_{form} for dodecane. As discussed in Section 3, these CS−MT model predictions of g_{form} were made using eq 9, with g_{dehydr} and g_{hydr} computed using eqs 11 and 12, respectively. A total of five bulk water and five aggregate simulations were required to generate these results. The second column of predictions for the free energy of formation (reported as $g_{form,max}$ in Table A2) were computed using the maximum value of $g_{tr,CS-MT}$ for dodecane given in Table A1 and using eq A2. As a result, only one bulk water and one aggregate simulation were required to generate these results. The maximum value of $g_{tr,CS-MT}$ obtained using the CS−MT model in Table A1 corresponds to the small spherical dodecane aggregate ($-21.88k_BT$). The third column of predictions for the free energy of formation (reported as $g_{form,min}$ in Table A2) were computed using the minimum value of $g_{tr,CS-MT}$ for dodecane given in Table A1 and using eq A2. Like the $g_{form,max}$ predictions, only one bulk water and one aggregate simulation were required to generate the $g_{form,min}$ predictions. As shown in Table A1, the minimum value of $g_{tr,CS-MT}$ was obtained on the basis of simulation results for the dodecane slab ($-22.06k_BT$). The absolute values of the percent differences between $g_{form,max}$ and $g_{form,min}$ (reported as $\|\Delta$ (%)$\|$) are also reported in Table A2.

The level of agreement between g_{form}, $g_{form,max}$, and $g_{form,min}$ is very high. The average value of $\|\Delta$ (%) $\|$ for the five aggregate geometries considered in Table A2 is only 1.23%. In addition, the reduction in computational cost associated with the use of eq A2 is very significant. Accordingly, we will use eq A2 in article 2 of this series to predict (i) optimal micelle shapes and sizes, (ii) the corresponding values of g_{form}, and (iii) CMCs of nonionic surfactants.

Appendix B: Mathematical Demonstration of the Equivalence of the CS−MT Model and the Traditional MT Model for Completely Hydrophobic Solutes

The CS−MT model and the traditional MT model for the hydrophobic effect are equivalent for completely hydrophobic solutes such as the oil molecules considered in this article. In this appendix, this equivalence will be demonstrated mathematically. For oil molecules, g_{dehydr} and g_{hydr} may be combined to obtain an expression analogous to the expression used to calculate g_{tr} and g_{int} in the traditional MT modeling approach. Recall that

$$g_{dehydr} = \sum_{i=1}^{n_{hydr}} (1 - f_i) g_{tr_i} \tag{B1}$$

and

$$g_{hydr} = \sum_{i=1}^{n_{core}} SASA_i f_i \left(\frac{\sigma A_{core}}{SASA_{core}} + \frac{g_{tr_i}}{SASA_i} \right) \tag{B2}$$

as introduced and discussed in Section 3 (see eqs 11 and 12). For oil molecules, the sum of these two free energies represents the total contribution to g_{form} due to the hydrophobic effect (HE). We will refer to this free-energy contribution as g_{HE}. The CS−MT model for g_{HE} may be expressed and manipulated as follows:

$$g_{HE} = \sum_{i=1}^{n_{hydr}} (1 - f_i) g_{tr_i} + \sum_{i=1}^{n_{core}} SASA_i f_i \left(\frac{\sigma A_{core}}{SASA_{core}} + \frac{g_{tr_i}}{SASA_i} \right) \tag{B3}$$

or

$$g_{HE} = \sum_{i=1}^{n_{hydr}} g_{tr_i} - \sum_{i=1}^{n_{hydr}} f_i g_{tr_i} + \frac{\sigma A_{core}}{SASA_{core}} \sum_{i=1}^{n_{core}} SASA_i f_i + \sum_{i=1}^{n_{core}} f_i g_{tr_i} \tag{B4}$$

For the oil molecules considered in this article, $n_{hydr} = n_{core}$ because every hydrophobic group in each oil molecule is incorporated into the aggregate core upon aggregate self-assembly. As a result, the second term and the last term on the right-hand side of eq B4 cancel out. We note that this is not the case for amphiphilic solutes such as surfactants. For oil molecules, eq B4 reduces to

$$g_{HE} = \sum_{i=1}^{n_{core}} g_{tr_i} + \frac{\sigma A_{core}}{SASA_{core}} \sum_{i=1}^{n_{core}} SASA_i f_i \tag{B5}$$

or to

$$g_{HE} = \sum_{i=1}^{n_{core}} g_{tr_i} + \sigma_{core} \langle SASA^{hydr} \rangle \tag{B6}$$

where $\sigma_{core} = \sigma A_{core}/SASA_{core}$ was introduced in Section 3.4 and represents the free energy of the hydrophobic core−water interface on a per unit SASA basis. Note that in eq B6 we have replaced the term $\sum_{i=1}^{n_{core}} SASA_i f_i$ with $\langle SASA^{hydr} \rangle$ to emphasize that the average of the product of $SASA_i$ and the value of f_i obtained through computer simulations should be equal to the average value of SASA that is exposed to hydrating contacts. For $\sum_{i=1}^{n_{core}} SASA_i f_i$ to be equal to $\langle SASA^{hydr} \rangle$ and for the CS−MT model to yield valid results, one requires that (i) each f_i value must be defined so that it conveys information about only the local environment of group i (i.e., the environment immediately surrounding the solvent accessible surface of group i)

and (ii) the approach used to determine $SASA_i$ and $SASA_{core}$ must give physically realistic results, where the same solute probe size must be used to quantify both SASA values. The close agreement between the CS−MT and the traditional MT modeling results obtained in this article demonstrates that both conditions (i) and (ii) above are satisfied for the analysis approach that we have presented.

The expression for g_{HE} in eq B6 in the context of the CS−MT model is directly analogous to the g_{HE} expression in the traditional MT model of the hydrophobic effect for aggregates containing solutes that are completely hydrophobic (i.e., aggregates where no solute heads are present):

$$g_{HE} = g_{tr} + g_{int} \tag{B7}$$

or

$$g_{HE} = \sum_{i=1}^{n_{core}} g_{tr_i} + \sigma a \tag{B8}$$

where σ is the interfacial tension of the aggregate core−water interface and a is the area of the hydrophobic core per solute molecule as computed geometrically based on the volume of the aggregate under the assumption of a perfectly smooth aggregate surface (see Section 3.4). Comparison of eqs B6 and B8 shows that, in the absence of solute heads, the CS−MT and the traditional MT models both include a term that quantifies the free-energy contribution associated with transferring each hydrophobic group from bulk water to a bulk phase of group i ($\sum_{i=1}^{n_{core}} g_{tr_i}$), as well as a term that quantifies the free-energy contribution associated with the formation of a hydrophobic core−water interface. In the CS−MT model, this free-energy contribution is computed using SASA and σ_{core}; in the traditional MT model, it is computed using a and σ.

The CS−MT model and the traditional MT model are not equivalent, however, for amphiphilic solutes possessing a head and a tail. For such solutes, n_{hydr} is not equal to n_{core}, and the CS−MT and traditional MT estimates for the free-energy contribution associated with transferring each solute from the bulk water state to the aggregate state will differ. In the CS−MT model, each of the hydrophobic groups in a solute contributes its transfer free energy to g_{HE} to the extent that it is dehydrated, whereas in the traditional MT model only the hydrophobic groups in the surfactant tail contribute their transfer free energies to g_{HE}. The CS−MT modeling approach avoids the "all-or-nothing" approximation implicit in the traditional MT modeling approach shown in eq B8, in which groups in the head do not contribute at all to g_{tr} and groups in the tail contribute fully their individual transfer free energies.

CS−MT and traditional MT model estimates of the free-energy contribution associated with the formation of the aggregate core−water interface are also not equivalent for amphiphilic solutes, because the two models account differently for the presence of the solute heads at the aggregate core−water interface. For aggregates containing amphiphilic solutes, the CS−MT modeling approach provides a way to directly calculate the hydrated SASA for each solute molecule, even when the solute heads are present, including using this SASA to compute the interfacial free energy. In Section 4.5, we argued that, when the CS−MT model is used to model amphiphilic solutes, both contacts with water and with hydrogen-bonding atoms in the solute head should be counted as hydrating contacts in computing each f_i value. On the other hand, when using the traditional MT model, an assumption must be made about the extent to which the solute heads shield the aggregate core−water interface from water contacts. The area screened by a surfactant head is traditionally approximated as being equal to the cross-sectional area of a single carbon−carbon bond or approximately 21 Å^2.[14] The validity of this assumption will be discussed in greater detail in article 2 of this series. Another difference between the CS−MT model and the traditional MT model estimates of the free-energy contribution associated with the formation of the aggregate core−water interface is that the CS−MT modeling approach permits estimation of the ensemble average of this free-energy contribution. This may be implemented in the CS−MT modeling approach by analyzing micelle microstructure and hydration at every snapshot in the molecular dynamics trajectory and determining the hydrophobic groups in each amphiphilic solute that are part of the aggregate core at each instant in time. For solute groups that are part of the aggregate core, g_{hydr} is calculated using eq B3. Determining the hydrophobic groups in each solute that are part of the aggregate core at any given time may be done in a number of ways, including comparing each group's f_i value to a cutoff value of f used to identify groups that are part of the aggregate core or determining each group's position relative to the aggregate core−water interface. An estimate of g_{HE} can then be made by averaging the g_{hydr} values obtained over the course of the simulation run. In this manner, the CS−MT model avoids the all-or-nothing approximation implicit in the traditional MT modeling approach to compute g_{int}, in which head and tail assignments must be made, the aggregate core is assumed to contain only tails, and eq B8 is invoked. A final difference between the way in which the free-energy contribution associated with forming the aggregate core−water interface is estimated in the CS−MT and in the traditional MT modeling approaches is that, by using the results from computer simulations, the CS−MT modeling approach enables estimation of the free-energy contribution associated with the formation of the aggregate core−water interface over an ensemble average of many different physically realistic micellar configurations, rather than of the three idealized, static configurations (a perfect sphere, a perfect cylinder, or a perfect bilayer) that can be successfully modeled in the traditional MT modeling approach.[14] The manner in which the traditional MT modeling results for these three idealized geometries are combined to model (i) finite cylinders with hemispherical endcaps and (ii) finite disk-like micelles is discussed in detail in ref 32.

References and Notes

(1) Elworthy, P. H.; Florence, A. T.; Macfarlane, C. B. *Solubilization by Surface Active Agents;* Chapman and Hall: London, 1968.

(2) Dunaway, C. S.; Christian, S. D.; Scamehorn, J. F. *Solubilization in Surfactant Aggregates. Surfactant Science Series 55*; Marcel Dekker: New York, 1995.

(3) Nagarajan, R. *Curr. Opin. Colloid Interface Sci.* **1996**, *1*, 391−401.

(4) Nagarajan, R. *Curr. Opin. Colloid Interface Sci.* **1997**, *2*, 282−293.

(5) Srinivasan, V. Thesis. *Theoretical Modeling of Micellization and Solubilization in Ionic Surfactant Systems*; Massachusetts Institute of Technology, 2003, and references cited therein.

(6) Tanford, C. *The Hydrophobic Effect: Formation of Micelles and Biological Membranes*; John Wiley and Sons: New York, 1991.

(7) Israelachvili, J. N. *Intermolecular and Surface Forces*, 2nd ed.; Academic Press: New York, 1991.

(8) Nagarajan, R.; Ruckenstein, E. *Langmuir* **1991**, *7*, 2934−2969.

(9) Shiloach, A.; Blankschtein, D. *Langmuir* **1998**, *14*, 1618−1636, and references cited therein.

(10) Gunnarsson, G.; Jonsson, B.; Wennerstrom, H. *J. Phys. Chem.* **1980**, *84*, 3114−3121.

(11) Jonsson, B.; Wennerstrom, H. *J. Colloid Interface Sci.* **1981**, *80*, 482−496.

(12) Evans, D. F.; Mitchell, D. J.; Ninham, B. W. *J. Phys. Chem.* **1984**, *88*, 6344−6348.
(13) Hayter, J. B. *Langmuir* **1992**, *8*, 2873−2876.
(14) Puvvada, S.; Blankschtein, D. *J. Chem. Phys.* **1990**, *92*, 3710−3724, and references cited therein.
(15) Srinivasan, V.; Blankschtein, D. *Langmuir* **2003**, *19*, 9932−9945, and references cited therein.
(16) Srinivasan, V.; Blankschtein, D. *Langmuir* **2003**, *19*, 9946−9961.
(17) Reif, I.; Mulqueen, M.; Blankschtein, D. *Langmuir* **2001**, *17*, 5801−5812.
(18) Shiloach, A.; Blankschtein, D. *Langmuir* **1998**, *14*, 7166−7182.
(19) Shiloach, A.; Blankschtein, D. *Langmuir* **1998**, *14*, 4105−4114.
(20) Shiloach, A.; Blankschtein, D. *Langmuir* **1997**, *13*, 3968−3981.
(21) Zoeller, N.; Lue, L.; Blankschtein, D. *Langmuir* **1997**, *13*, 5258−5275.
(22) Goldsipe, A.; Blankschtein, D. *Langmuir* **2005**, *22*, 9850−9865, and references cited therein.
(23) Molecular Modeling Pro, version 3.2; ChemSW, Inc.: Fairfield, CA, 2003.
(24) Stephenson, B. C.; Beers, K.; Blankschtein, D. *Langmuir* **2006**, *22*, 1500−1513, and references cited therein.
(25) Stephenson, B. C.; Rangel-Yagui, C. O.; Pessoa, A.; Tavares, L. C.; Beers, K.; Blankschtein, D. *Langmuir* **2006**, *22*, 1514−1525.
(26) Stephenson, B. C.; Goldsipe, A.; Beers, K. J.; Blankschtein, D. *J. Phys. Chem. B* **2007**, *111*, 1045−1062.
(27) Stephenson, B. C.; Beers, K. J.; Blankschtein, D. *J. Phys. Chem. B* **2007**, *111*, 1063−1075.
(28) Stephenson, B. C.; Mendenhall, J.; Beers, K. J.; Blankschtein, D. *J. Phys. Chem. B*, in preparation.
(29) Ben-Shaul, A.; Szleifer, I. *J. Chem. Phys.* **1985**, *83*, 3597−3611.
(30) Szleifer, I.; Ben-Shaul, A.; Gelbart, W. M. *J. Chem. Phys.* **1985**, *83*, 3612−3620.
(31) Szleifer, I.; Ben-Shaul, A.; Gelbart, W. M. *J. Chem. Phys.* **1987**, *86*, 7094−7109.
(32) Srinivasan, V.; Blankschtein, D. *Langmuir* **2005**, *21*, 1647−1660.
(33) Puvvada, S.; Blankschtein, D. *J. Phys. Chem.* **1992**, *96*, 5579−5592.
(34) Maibaum, L.; Dinner, A. R.; Chandler, D. *J. Phys. Chem. B* **2004**, *108*, 6778−6781.
(35) Smith, R.; Tanford, C. *Proc. Natl. Acad. Sci.* **1973**, *70*, 289−293.
(36) Carale, T. R.; Pham, Q. T.; Blankschtein, D. *Langmuir* **1994**, *10*, 109−121.
(37) Gibbs, J. W. *The Scientific Papers of J.W. Gibbs*; Dover: New York, 1961; Vol. 1.
(38) Koenig, F. O. *J. Chem. Phys.* **1950**, *18*, 449.
(39) Buff, F. P. *J. Chem. Phys.* **1951**, *19*, 1591.
(40) Tolman, R. C. *J. Chem. Phys.* **1948**, *16*, 758.
(41) Aveyard, R.; Briscoe, B. J.; Chapman, J. *J. Chem. Soc., Faraday Trans. 1* **1972**, *68*, 10.
(42) Lum, K.; Chandler, D.; Weeks, J. D. *J. Phys. Chem. B* **1999**, *103*, 4570−4577.
(43) Chandler, D. *Nature* **2005**, *437*, 640−647.
(44) Reynolds, J. A.; Gilbert, D. B.; Tanford, C. *Proc. Natl. Acad. Sci.* **1974**, *71*, 2925−2927.
(45) Hermann, R. B. *J. Phys. Chem.* **1972**, *76*, 2754−2759.
(46) Pool, R.; Bolhuis, P. G. *J. Phys. Chem. B* **2005**, *109*, 6650−6657.
(47) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.
(48) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269−6271.
(49) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(50) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(51) Bader, J. S.; Chandler, D. *J. Phys. Chem.* **1992**, *96*, 6423−6427.
(52) Shirts, M. R.; Pitera, J. W.; Swope, W. C.; Pande, V. S. *J. Chem. Phys.* **2003**, *119*, 5740−5761.
(53) Shirts, M. R.; Pande, V. S. *J. Chem. Phys.* **2005**, *122*, 134508.
(54) Shirts, M. R. Thesis. *Calculating Precise and Accurate Free Energies in Biomolecular Systems*; Stanford University, 2005, and references cited therein.
(55) van der Spoel, D.; Lindahl, E.; Hess, B.; van Buuren, A. R.; Apol, E.; Meulenhoff, P. J.; Tieleman, D. P.; Sijbers, A. L. T. M.; Feenstra, K. A.; van Drunen, R.; Berendsen, H. J. C. *Gromacs User Manual version 3.2*; www.gromacs.org:2004.
(56) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1997**, *23*, 327−341.
(57) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. *Comput. Phys. Commun.* **1995**, *91*, 43−56.
(58) Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Model.* **2001**, *7*, 306−317.
(59) Jedlovszky, P. *J. Chem. Phys.* **2003**, *119*, 1731−1740.
(60) Flyvbjerg, H.; Petersen, H. G. *J. Chem. Phys.* **1989**, *91*, 461−466.
(61) Hess, B. Thesis. *Stochastic Concepts in Molecular Simulation*; Rijksuniversiteit Groningen, 1999.
(62) Hess, B. *J. Chem. Phys.* **2002**, *116*, 209−217.
(63) Pearlman, D. A.; Kollman, P. A. *J. Chem. Phys.* **1989**, *91*, 7831−7839.

Quantifying the Hydrophobic Effect. 2. A Computer Simulation−Molecular-Thermodynamic Model for the Micellization of Nonionic Surfactants in Aqueous Solution

Brian C. Stephenson, Arthur Goldsipe, Kenneth J. Beers, and Daniel Blankschtein*

Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received: September 1, 2006; In Final Form: November 21, 2006

In this article, the validity and accuracy of the CS−MT model is evaluated by using it to model the micellization behavior of seven nonionic surfactants in aqueous solution. Detailed information about the changes in hydration that occur upon the self-assembly of the surfactants into micelles was obtained through molecular dynamics simulation and subsequently used to compute the hydrophobic driving force for micelle formation. This information has also been used to test, for the first time, approximations made in traditional molecular-thermodynamic modeling. In the CS−MT model, two separate free-energy contributions to the hydrophobic driving force are computed. The first contribution, g_{dehydr}, is the free-energy change associated with the dehydration of each surfactant group upon micelle formation. The second contribution, g_{hydr}, is the change in the hydration free energy of each surfactant group upon micelle formation. To enable the straightforward estimation of g_{dehydr} and g_{hydr} in the case of nonionic surfactants, a number of simplifying approximations were made. Although the CS−MT model can be used to predict a variety of micellar solution properties including the micelle shape, size, and composition, the critical micelle concentration (CMC) was selected for prediction and comparison with experimental CMC data because it depends exponentially on the free energy of micelle formation, and as such, it provides a stringent quantitative test with which to evaluate the predictive accuracy of the CS−MT model. Reasonable agreement between the CMCs predicted by the CS−MT model and the experimental CMCs was obtained for octyl glucoside (OG), dodecyl maltoside (DM), octyl sulfinyl ethanol (OSE), decyl methyl sulfoxide ($C_{10}SO$), decyl dimethyl phosphine oxide ($C_{10}PO$), and decanoyl-*n*-methylglucamide (MEGA-10). For five of these surfactants, the CMCs predicted using the CS−MT model were closer to the experimental CMCs than the CMCs predicted using the traditional molecular-thermodynamic (MT) model. In addition, CMCs predicted for mixtures of $C_{10}PO$ and $C_{10}SO$ using the CS−MT model were significantly closer to the experimental CMCs than those predicted using the traditional MT model. For dodecyl octa(ethylene oxide) ($C_{12}E_8$), the CMC predicted by the CS−MT model was not in good agreement with the experimental CMC and with the CMC predicted by the traditional MT model, because the simplifying approximations made to estimate g_{dehydr} and g_{hydr} in this case were not sufficiently accurate. Consequently, we recommend that these simplifying approximations only be used for nonionic surfactants possessing relatively small, non-polymeric heads. For MEGA-10, which is the most structurally complex of the seven nonionic surfactants modeled, the CMC predicted by the CS−MT model (6.55 mM) was found to be in much closer agreement with the experimental CMC (5 mM) than the CMC predicted by the traditional MT model (43.3 mM). Our results suggest that, for complex, small-head nonionic surfactants where it is difficult to accurately quantify the hydrophobic driving force for micelle formation using the traditional MT modeling approach, the CS−MT model is capable of making reasonable predictions of aqueous micellization behavior.

1. Introduction

In article 1 of this series,[1] we developed a computer simulation−molecular-thermodynamic (CS−MT) modeling approach to better understand and quantify the hydrophobic driving force for solute (surfactant and solubilizate) aggregate formation in aqueous solution. As discussed in article 1, a significant body of literature on traditional MT modeling has demonstrated its ability to model the micellization behavior of structurally simple surfactants with quantitative or semiquantitative accuracy.[2−7] In the traditional MT modeling approach, the free-energy change associated with the formation of the surfactant aggregate in aqueous solution is expressed as the sum of several free-energy contributions, all of which can be computed molecularly given the chemical structures of the various micellar components and the solution conditions. To date, traditional MT models of micellization and micellar solubilization have relied on relatively simple approximations for the micellar hydration states of the surfactants and the solubilizates. To extend the applicability of the traditional MT modeling approach to more chemically and structurally complex surfactants and solubilizates, there is a need to accurately estimate the hydration states of these solutes in the micellar state. The CS−MT model represents a novel approach to obtain and analyze this type of hydration data. With the above in mind, in this article, we use the CS−MT model to predict the micellization behavior of nonionic surfactants that are both simple and challenging to model using the traditional MT modeling approach.

1.1. Review of the CS−MT Model. In the CS−MT model, the free energy of aggregate formation, g_{form}, is computed as

* Corresponding author. Department of Chemical Engineering, Room 66-444, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139. Telephone: (617) 253-4594. Fax: (617) 252-1651. E-mail: dblank@mit.edu

10.1021/jp065697a CCC: $37.00
Published on Web 01/18/2007

the sum of the following six free-energy contributions:[1]

$$g_{\text{form}} = g_{\text{dehydr}} + g_{\text{hydr}} + g_{\text{pack}} + g_{\text{st}} + g_{\text{elec}} + g_{\text{ent}} \quad (1)$$

The physical origin of each of these free-energy contributions can be understood by representing the process of aggregate formation as a thermodynamic cycle consisting of three separate steps (see Figure 2 in article 1). Two of the free-energy contributions in eq 1, g_{dehydr} and g_{hydr}, reflect the hydrophobic free-energy change associated with aggregate formation, or the hydrophobic driving force for micelle formation. In the CS−MT modeling approach, both g_{dehydr} and g_{hydr} are computed using hydration data obtained from computer simulations. The remaining four free-energy contributions (g_{pack}, g_{st}, g_{elec}, and g_{ent}) in eq 1 are computed in the CS−MT model in the same manner as they are computed in the traditional MT modeling approach.[8] However, the way in which g_{pack} and g_{st} are computed could, in principle, be informed by the molecular dynamics simulation data. In article 1, we proposed and validated theoretical models to evaluate g_{dehydr} and g_{hydr}. The free-energy contribution, g_{dehydr}, is computed as follows:[1]

$$g_{\text{dehydr}} = \sum_{i=1}^{n_{\text{hydr}}} (1 - f_i) g_{\text{tr}_i} \quad (2)$$

where n_{hydr} is the total number of hydrophobic groups in the solute, f_i is the fractional hydration of group i, and g_{tr_i} is the free-energy change associated with transferring group i from the aqueous solution to a bulk phase of solute tails. In article 1, we justified computing f for each group i as follows:

$$f = \frac{\text{number of hydrating contacts in the aggregate}}{\text{number of hydrating contacts in bulk water}} \quad (3)$$

where a "hydrating contact" is defined as a contact with an atom that (i) forms hydrogen bonds or (ii) is capable of coordinate (dative covalent) bonding. On the basis of this definition, if a hydrophobic CH_2 group is in contact with the oxygen or hydrogen atoms of a water molecule, a positively charged or a negatively charged ion, or a hydrophilic group in the solute head that is capable of hydrogen bonding, then the contact is considered hydrating. In article 1, we also justified the use of a 0.3 nm cutoff distance to count the hydrating contacts that occur during MD simulation.

The free-energy contribution, g_{hydr}, is computed as follows:[1]

$$g_{\text{hydr}} = \sum_{i=1}^{n_{\text{core}}} \text{SASA}_i f_i \Delta g_{\text{wc}_i} \quad (4)$$

where n_{core} is the total number of hydrophobic groups in the solute that adsorb onto, or penetrate into, the aggregate core, SASA_i is the solvent accessible surface area of group i, and Δg_{wc_i} is defined as the difference in the free energy per unit of solvent accessible surface area associated with the hydration of group i in the micellar state and in the aqueous solution.

In article 1, the CS−MT model was used to calculate the free-energy change associated with the formation of aggregates of octane, dodecane, and hexadecane having various shapes (spheres, cylinders, and slabs) and sizes. In total, five different aggregate geometries were considered for each oil type. To compute g_{form}, f_i data was calculated using information on water contacts obtained by simulating a single oil molecule in bulk water and by simulating the same oil molecule in an oil aggregate. Values of g_{tr_i} were estimated for the CH_2 and the CH_3 groups in each oil molecule from aqueous solubility data of linear alkanes. Values of SASA_i for the CH_2 and the CH_3 groups were estimated using the double cubic lattice method as implemented in GROMACS and a solvent probe of radius 0.2 nm.[9] We also developed an approach to theoretically calculate Δg_{wc_i} for oil molecules, in which Δg_{wc_i} is calculated as the difference between two "microscopic interfacial tensions" or free energies per unit SASA. For oil molecules, Δg_{wc_i} does not depend on i, and is given by[1]

$$\Delta g_{\text{wc}} = \sigma_{\text{core}} - \sigma_{\text{bulk}} = \frac{\sigma A_{\text{core}}}{\text{SASA}_{\text{core}}} - \frac{g_{\text{tr}_i}}{\text{SASA}_i} \quad (5)$$

where σ_{core} is the microscopic "interfacial tension" (interfacial free energy per unit SASA) associated with the aggregate core−water interface, σ_{bulk} is the microscopic "interfacial tension" (interfacial free energy per unit SASA) associated with the group i (CH_2 or CH_3)−water interface in the aqueous solution, σ is the macroscopic interfacial tension of the aggregate core−water interface, A_{core} is the surface area of the hydrophobic aggregate core as computed geometrically based on the volume of the aggregate subject to the assumption of a perfectly smooth aggregate surface, and $\text{SASA}_{\text{core}}$ is the solvent accessible surface area of the hydrophobic aggregate core.

The ratio $A_{\text{core}}/\text{SASA}_{\text{core}}$ in eq 5 was estimated using the following correlation that was fitted on the basis of our computer simulation results for the various oil aggregates considered:[1]

$$\text{SASA}_{\text{core}}/A_{\text{core}} = 1.740 - 0.026 n_{\text{t}} + 0.078 C \quad (6)$$

where n_{t} is the total number of hydrophobic groups in the solute that are part of the hydrophobic aggregate core and C is the curvature of the micellar aggregate, which is defined as $2/l_{\text{c}}$ for spheres, $1/l_{\text{c}}$ for cylinders, and zero for planar interfaces, where l_{c} is the core-minor radius or planar half-width. In article 1, excellent agreement between the predictions of the CS−MT model and the traditional MT model was obtained for g_{form} for each of the 15 oil aggregates modeled, with an average absolute error of only 1.04% between the two modeling approaches. The very high level of agreement between the CS−MT and the traditional MT modeling results demonstrates the ability of the CS−MT model to quantify the hydrophobic effect for completely hydrophobic solutes, as well as to calculate g_{form} with a high degree of accuracy.

In Appendix A of article 1, we showed that, by combining elements of the CS−MT model and the traditional MT model, g_{form} can be computed as a function of aggregate shape and size after performing *only two computer simulations*: the first of the solute in a bulk water environment and the second of the same solute in an aggregate environment (where the aggregate can have arbitrary shape and size). Specifically, we showed that[1]

$$g_{\text{tr,CS-MT}} = g_{\text{dehydr}} + g_{\text{hydr}} - \hat{g}_{\text{int}} \quad (7)$$

where $g_{\text{tr,CS-MT}}$ is the transfer free-energy contribution computed using the CS−MT modeling approach and $\hat{g}_{\text{int}}$ is the traditional MT prediction for the interfacial free-energy contribution of the simulated micellar aggregate. The free energy of aggregate formation, g_{form}, for a micelle of a different shape and size than that at which the computer simulation data was gathered is then calculated using the following equation:[1]

$$g_{\text{form}} = g_{\text{tr,CS-MT}} + g_{\text{int}} + g_{\text{pack}} + g_{\text{st}} + g_{\text{elec}} + g_{\text{ent}} \quad (8)$$

In article 1, we demonstrated that consistent values of $g_{\text{tr,CS-MT}}$ based on hydration information obtained through

computer simulation of oil aggregates of different curvatures can be estimated using eq 7. Using the computed value of $g_{tr,CS-MT}$, we also demonstrated that highly accurate values of g_{form} could be obtained using eq 8 for each of the 15 oil aggregates considered.

For a micelle of the optimum shape, size, composition (in the case of mixed micelles), and degree of counterion binding (in the case of ionic surfactants), g_{form} has a minimum value, which we denote as g^*_{form}. By solving for g^*_{form}, the optimal aggregate shape (S^*), the optimal core-minor radius (l^*_c), the optimal composition (α^*), and the optimal degree of counterion binding (β^*), can be predicted. In addition, the CMC in mole fraction units is computed as follows:[10]

$$\mathrm{CMC} \approx \exp\left(\frac{g^*_{form}(S^*,l^*_c,\alpha^*,\beta^*)}{k_B T}\right) \quad (9)$$

where k_B is the Boltzmann constant and T is the absolute temperature.

1.2. Modeling Nonionic Surfactant Micellization. In this article, we will use the CS−MT modeling approach introduced in article 1 to model the micellization behavior of nonionic surfactants. Although the CS−MT model enables the prediction of a wide range of solution properties, the CMC has been selected for prediction and comparison with the experimental CMC data because the CMC depends exponentially on g_{form}, and as such, it provides a stringent test with which to evaluate the predictive accuracy of the CS−MT model. We have selected the following seven nonionic surfactants in order to test and validate the CS−MT model: octyl glucoside (OG), dodecyl maltoside (DM), octyl sulfinyl ethanol (OSE), decyl methyl sulfoxide ($C_{10}SO$), decyl dimethyl phosphine oxide ($C_{10}PO$), dodecyl octa(ethylene oxide) ($C_{12}E_8$), and decanoyl-*n*-methylglucamide (MEGA-10). These seven nonionic surfactants have varying degrees of structural and chemical complexity and, as such, have allowed us to thoroughly gauge the validity and predictive accuracy of the CS−MT modeling approach.

In order to use the CS−MT modeling approach in the case of the nonionic surfactants considered here, we have made three approximations to calculate g_{dehydr} and g_{hydr} in a relatively simple manner. The first approximation involves the way in which we estimate g_{tr_i} to enable the evaluation of g_{dehydr} using eq 2. The second approximation involves introducing an approach to determine which surfactant groups are adsorbed onto, or incorporated within, the micelle core to enable the evaluation of g_{hydr} using eq 4. The third approximation involves using the theoretical model for Δg_{wc}, given in eqs 5 and 6, which was developed for oil molecules, in the case of nonionic surfactants (which are amphiphilic solutes). The validity of these three approximations will be discussed in Section 4.

In addition to determining the validity of the CS−MT model in the case of nonionic surfactants, we will use the detailed hydration information obtained through computer simulation of nonionic surfactant micelles to quantitatively evaluate several of the approximations underlying the traditional MT modeling approach. Specifically, we will evaluate (i) the accuracy of computing the transfer free-energy contribution, g_{tr}, using the head and tail approximations made in the context of the traditional MT modeling approach,[1] and (ii) the extent to which the surfactant heads shield the micelle hydrophobic core from hydrating contacts.

The remainder of this article is organized as follows. Section 2 describes the computer simulation approach that we have used, including an overview of the modeling approach (Section 2.1), the simulation methods and parameters (Section 2.2), and a description of how each system has been prepared and equilibrated (Section 2.3). The data analysis method used to analyze the molecular dynamics trajectories is described in Section 2.4. In Sections 3.1−3.7, computer simulation results are presented for each of the seven nonionic surfactants modeled in this article. In Section 3.8, the accuracy of several approximations made in the traditional MT modeling approach is determined on the basis of the computer simulation results. In Section 4, the CS−MT model and the traditional MT model are used to model the micellization behavior of each of the seven nonionic surfactants considered. Finally, concluding remarks are presented in Section 5.

2. Molecular Dynamics Simulations

2.1. Modeling Approach. To quantify the hydrophobic driving force associated with the formation of nonionic surfactant micelles, we have used atomistic-level computer simulations to determine the change in hydration for each atom (or group of atoms, such as a CH_2 group) upon being transferred from the aqueous solution to the aggregate environment. As described in article 1, this is accomplished by performing two simulations. The first simulation is of a single nonionic surfactant in a simulation cell of water, which we will refer to hereafter as the "bulk water" simulation. The second simulation is of the same nonionic surfactant in a micellar environment, which we will refer to hereafter as the "aggregate" simulation. Each aggregate simulation was prepared by preforming a nonionic surfactant micelle at an arbitrary aggregation number. The nonionic surfactant micelle was simulated for 10 to 15 ns, which provides sufficient time for the surfactant molecules within the micelle to rearrange and come to local equilibrium, but does not provide sufficient time for the surfactant molecules to exit the aggregate environment and enter the aqueous solution. As a result, the computer simulation results do not permit direct prediction of the optimal micelle shape and size that would be observed experimentally. However, as shown in Appendix A of article 1, by using the CS−MT modeling approach, obtaining information about the hydration state of a micelle of a single shape and size is sufficient to allow prediction of the *optimal* micelle shape and size.

2.2. Simulation Methods and Parameters. The simulation methods and parameters used here are identical to those described in article 1, where we provided a detailed description of the simulation methodology.[1,14−19] Each of the nonionic surfactants was modeled using the fully atomistic OPLS-AA force field,[11] and water was modeled using the simple extended point-charge (SPC/E) model. For the nonionic surfactants OG, DM, and $C_{12}E_8$, atomic charges were assigned on the basis of the default atomic charge values recommended in OPLS-AA. However, because the OSE, $C_{10}SO$, $C_{10}PO$, and MEGA-10 surfactant head structures did not have suggested charges in the OPLS-AA force field, we estimated the atomic charges for these heads using the CHelpG algorithm (as implemented in Gaussian 98), in which atomic charges are assigned to fit electrostatic potentials at a number of points on the van der Waals surface.[12] We note that CHelpG was *not* used to assign atomic charges for the hydrophobic tails of these four surfactants. In a separate study, we determined that assigning atomic charges to the CH_2 and the CH_3 groups in a linear alkyl chain using CHelpG yields simulation results that are less physically realistic than those obtained by assigning atomic charges with the recommended OPLS-AA charges (results not shown).

In two recent publications, we investigated the sensitivity of the head and tail assignments obtained through computer

TABLE 1: Overview of the MD Simulations[a]

surfactant	number of surfactant molecules	number of water molecules	total number of atoms
OG	29	3695	12477
DM	45	4283	16494
OSE	25	3750	12125
$C_{12}E_8$	41	15256	49663
$C_{10}PO$	50	6708	22174
$C_{10}SO$	50	3510	12380
MEGA−10	42	3351	12531

[a] The number of surfactant and water molecules and the total number of atoms corresponding to each of the seven simulated nonionic surfactant micelles are listed.

simulation to the method used to assign atomic charges.[13,14] In general, we found that the results are sensitive to the atomic charges used and that the charge assignments recommended within the OPLS-AA force field yield more reasonable results than those obtained using the CHelpG algorithm. However, if a specific surfactant or solubilizate does not have suggested charges in the OPLS-AA force field, we found that applying the CHelpG approach to determine charges yields reasonably accurate results.

2.3. System Preparation and Equilibration. *2.3.1. Bulk Water Simulation.* The bulk water simulation for each of the nonionic surfactants considered was initialized by placing a single surfactant molecule in a simulation cell and surrounding it with water molecules. The simulation cell was selected to be sufficiently large that there was always at least 2.0 nm of water separating the surfactant molecule from its periodic image, where this simulation cell size was justified in article 1.[1] After brief equilibration under *NPT* conditions until the system volume had stabilized, a 2 to 5 ns data-gathering simulation was carried out.

2.3.2. Aggregate Simulation. The method used to carry out each of the surfactant aggregate simulations was more complex. Each nonionic surfactant micelle was preformed as a spherical aggregate by placing a number of surfactant molecules in close proximity with each surfactant head oriented radially outward from the micelle center. The surfactant molecules were placed such that the surfactant heads were approximately uniformly spaced at the micelle surface. Next, sufficient water molecules were added around each micelle such that it was separated by at least 2 nm from its periodic image. A relatively large simulation cell size was required for $C_{12}E_8$ because of the large size of its polymeric E_8 head. The number of surfactant and water molecules and the total number of atoms included in the simulation cell for each nonionic surfactant micelle are listed in Table 1.

Selection of the Simulation Geometry. At this point, it is worth discussing why spherical, rather than cylindrical or bilayer, micelles were selected for simulation. As shown in Appendix A of article 1, any aggregation number and aggregate geometry (whether spherical, cylindrical, or planar) may be used to obtain hydration information for input to the CS−MT modeling approach.[1] An infinite cylinder or bilayer can be modeled in a computationally efficient manner by simulating only a small cross section of the cylinder or bilayer. We have selected spherical geometries for simulation, however, because carrying out physically realistic cylindrical and bilayer simulations requires that each surfactant molecule has a physically realistic area available to it at the micelle core−water interface (which we will refer to hereafter as *a*) for the simulated micelles to remain stable. The equilibrium area per surfactant head in a micelle results from a complex interplay of forces (including steric, electrostatic, van der Waals, and hydrogen-bonding interactions). If *a* for a preformed spherical micelle is too small, then the micelle will simply become somewhat ellipsoidal during simulation. On the other hand, if *a* for a preformed cylindrical or bilayer micelle is too small, the simulation cell dimensions must be allowed to expand or the micelle will buckle during the simulation and may break apart. If *a* is much larger than the experimental value, a spherical micelle will remain stable on the simulation timescales involved in CS−MT modeling, but a cylindrical micelle or bilayer may break up to form smaller spherical or ellipsoidal aggregates during simulation. In addition to these stability concerns, we note that if *a* is very different from the experimental value, we expect that eqs 7 and 8, in which $g_{tr,CS-MT}$ is assumed to be constant and not to be a function of micelle shape and size, may not be valid.

An appropriate value of *a* for each surfactant molecule in a cylindrical or planar aggregate can be determined through computer simulation by performing constant volume simulations of a number of cylindrical or planar aggregates preformed with different *a* values and using the computer simulation results to identify the value of *a* that yields the minimum energy of interface formation. Such an approach was implemented recently by Jang et al. in determining an appropriate *a* value for simulation of Newton black films.[20] An alternative approach to ensure that cylindrical or bilayer micelles are simulated with a reasonable value of *a* for each surfactant head is to preform the micelle at an arbitrary value of *a* per surfactant molecule but then allow the simulation cell dimensions to change during simulation subject to physically realistic boundary conditions. Unfortunately, such boundary conditions are difficult to determine. The appropriate boundary condition to use parallel to the axis of a cylindrical micelle, or parallel to the surface of a bilayer micelle, is a surface tension that provides a post-equilibration value of *a* that is similar to the one that would be observed experimentally. Particularly in the case of bilayer simulations, one might assume that macroscopic surface tension data could be used to infer the appropriate surface tension value for use during simulation. However, it is known that macroscopically observed surface tensions are difficult to predict accurately from a microscopic simulation.[21,22] A major reason for this is that long (micron) wavelength undulations are not included in nanometer-scale simulation results. Other researchers have commented on this limitation and used it to justify applying nonzero surface tensions in flaccid lipid bilayer simulations where the appropriate macroscopic surface tension is arguably zero.[22] Accordingly, a macroscopic surface tension is not likely to be appropriate for use as a boundary condition during cylindrical or bilayer micelle simulation.

Clearly, the simulation of cylindrical or bilayer aggregates introduces complications that are not present during simulation of spherical aggregates. With these complications in mind, we chose to preform each surfactant micelle in a spherical geometry in aqueous solution. Each of the nonionic surfactant micelles was constructed with an aggregation number sufficiently small to ensure that it would exist as a spherical aggregate during simulation. For several surfactants (OG, DM, $C_{12}E_8$, and MEGA-10), this was accomplished by estimating the expected aggregation number of a spherical micelle given the head area and tail volume of each surfactant molecule.[23] For other surfactants (OSE, $C_{10}PO$, and $C_{10}SO$), the surfactant head area is sufficiently small that we would expect them to form cylindrical micelles with a potentially large aggregation number.[23] Therefore, for each of these three surfactants, spherical

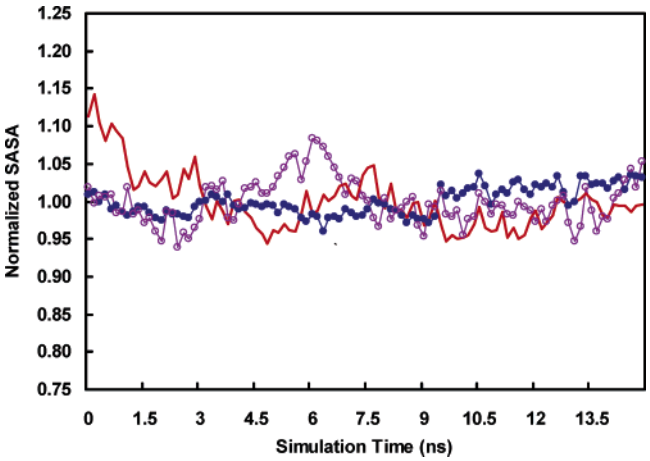


Figure 1. Solvent accessible surface area (SASA) normalized by the average value of SASA as a function of simulation time for micelles of three representative nonionic surfactants: dodecyl maltoside (DM, −●−), dodecyl octa(ethylene oxide) ($C_{12}E_8$, −○−), and decanoyl-*n*-methylglucamide (MEGA-10, —).

micelles were preformed with an aggregation number that was selected arbitrarily.

Micelle Equilibration. After preforming each spherical micelle, an energy minimization was conducted to remove close contacts. Next, an extended equilibration run under *NPT* conditions was conducted for 10 ns. Results by other researchers when conducting atomistic-level simulations of micelles in aqueous solution suggest that a simulation time of 10 ns should be more than adequate to equilibrate a spherical micelle.[24] One measure of equilibration for micellar systems is whether or not each group in a surfactant molecule has come to an equilibrium distance from the micelle center-of-mass. Bruce et al. have reported that sodium counterions are the slowest component of an SDS surfactant−water system to come to an equilibrium distance from the SDS micelle center-of-mass, taking only about 1 ns to equilibrate.[24] For the nonionic surfactants considered here, no counterions were present. Equilibration was confirmed from our simulation results by monitoring the total potential energy (which became stable during a small fraction of the total simulation time) and the solvent accessible surface area (SASA) of the micelle, where SASA was computed using the double cubic lattice method as implemented in GROMACS. The solvent accessible surface was traced out by a probe sphere of radius 0.2 nm (as justified in article 1)[1] that was rolled around each molecule within the aggregate to identify the solvent accessible region.[9] We consider SASA to be the most important metric to measure equilibration because this property is directly proportional to the degree of hydration of the micelle, and obtaining accurate hydration information is the central objective of our computer simulations. Plots of the SASA profiles for three representative nonionic surfactants (DM, $C_{12}E_8$, and MEGA-10) during equilibration are shown in Figure 1. The SASA values reported in Figure 1 for each surfactant have been normalized by the average SASA value for that surfactant to facilitate comparison of the results. The lack of noticeable drift in SASA toward the end of the 10 ns equilibration simulation run confirms that water contact data gathered during the subsequent 5 ns of data gathering should be representative of the hydration state of the micelle in its equilibrium configuration. Plots of the normalized SASA values over the course of the 5 ns data-gathering simulation runs for each surfactant are presented in the Supporting Information.

Snapshots of the post-equilibration configurations of each simulated nonionic micelle are shown in Figure 2. Each surfactant molecule is depicted using the van der Waals radius of each atom. For clarity, the water molecules are not shown.

2.4. Data Analysis Method. To quantify the degree of hydration of each atom (or group of atoms) in the surfactant molecule during the bulk water simulation, the number of contacts with hydrogen-bonding or with coordinate (dative covalent) bonding atoms per time step experienced by each atom was counted over the course of a simulation run, as justified in article 1.[1] For the nonionic surfactants considered in this article, contacts with both water atoms and with hydrogen-bonding surfactant headgroups have been counted as contributing to hydration. In analyzing our simulation data, a contact was defined as two atoms separated by less than 0.3 nm (the "cutoff" distance) at any time during the simulation. The average number of contacts is directly proportional to the average number of hydrogen or coordinate bonding atoms located within the cutoff distance. The degree of hydration of the surfactant molecules during the aggregate simulations was quantified in the same manner and with the same 0.3 nm cutoff distance used in analyzing the results of the bulk water simulation. From the contacts data obtained in this manner, we computed f, the fractional degree of hydration of each surfactant atom (or group of atoms), which is the key computer simulation input to the CS−MT model (see eq 3).

Although a cutoff of 0.3 nm was used to determine the hydration data for CS−MT modeling (which includes only f values for CH, CH_2, and CH_3 groups), a cutoff of 0.5 nm was used to generate the hydration plots presented in Section 3. Using a larger cutoff when generating the hydration plots improved the statistics of the f values obtained for several of the large atoms present in the surfactant heads (including nitrogen, sulfur, phosphorus, and oxygen).

An estimate of the standard error in f for each group of atoms in the surfactant molecule was made through block averaging, a useful approach to analyze correlated data.[25−27] A detailed discussion of this error analysis approach was presented in article 1.[1] Data-gathering simulation runs for each surfactant molecule in the bulk water and in the aggregate states were conducted for sufficient time to ensure that the uncertainty in each calculated value of f was small—typically less than 5%.

3. Simulation Results and Discussion

In this section, computer simulation results for fractional hydration are presented for each of the seven nonionic surfactants considered in this article. In addition to the f values computed using eq 3, two other fractional hydration values were computed: (i) f values in which the only contacts in the aggregate state that were counted as hydrating were water contacts (denoted as f_{water}) and (ii) f values in which the only contacts in the aggregate state that were counted as hydrating were contacts with hydrogen-bonding groups in the surfactant heads (denoted as f_{head}), where

$$f = f_{water} + f_{head} \tag{10}$$

We have computed values of f_{water} and f_{head} to gain insight into (i) the extent to which contacts with hydrophilic groups in the head contribute to the hydration of hydrophobic atoms in each surfactant head and tail and (ii) to determine whether f or f_{water} values are most appropriate to use in the CS−MT modeling approach. We can evaluate (i) on the basis of the hydration results for each surfactant presented in this section (see Figures 3−9). We will discuss (ii) in greater detail in Section 4.10, where

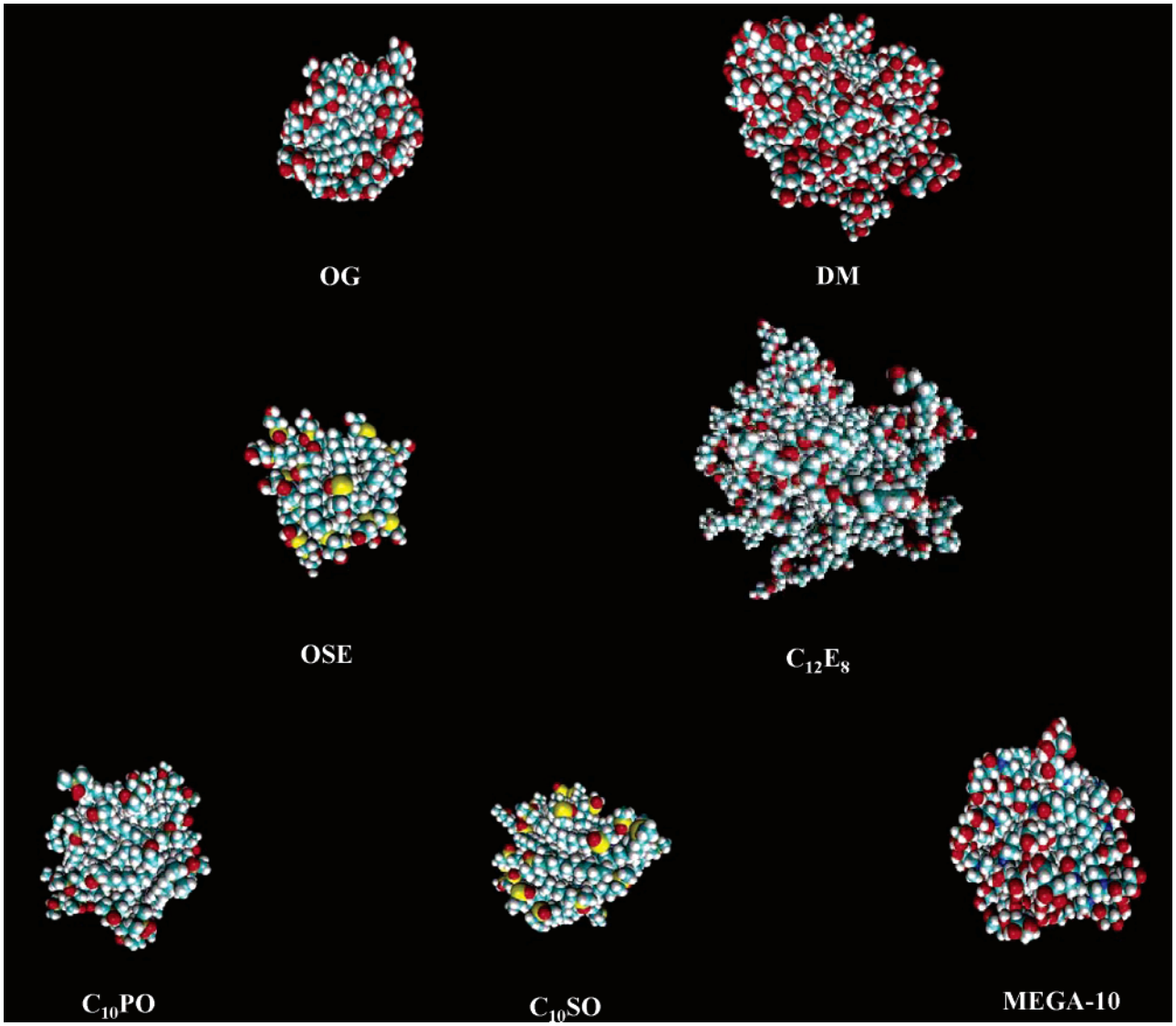


Figure 2. Snapshots of the post-equilibration structures of the seven simulated micelles considered here. The water molecules are not shown for clarity.

CS−MT modeling results obtained using f and f_{water} values will be compared.

3.1. Octyl Glucoside (OG). The fractional degree of hydration of OG is plotted as a function of group number in Figure 3. Three different fractional hydration profiles (f, f_{water}, and f_{head}) are shown in the figure. Hydrophilic and hydrophobic groups that are considered to be part of the OG head in traditional MT modeling (groups 1−12) have f values (see the −♦− results) that are much larger than the f values of groups in the OG tail (groups 13−20). However, to some extent, each group in the OG head is partially dehydrated and has an f value that is less than 1.0. The average f value of the groups in the OG head is 0.69. Clearly, the approximation made in traditional MT modeling that the surfactant head remains completely hydrated is not very accurate, although we note that for simple surfactants, the traditional MT modeling approach yields quantitatively, or semiquantitatively, accurate predictions of the micellization behavior.[2,3,8,10]

The f results in Figure 3 show that the degree of dehydration of the groups in the surfactant head is a function of their distance from the surfactant tail. For example, oxygen atom 12 (closest to the tail) has an f value of 0.4, and oxygen atom 10 has an f value of 0.7. Similarly, the degree of dehydration of the groups in the surfactant tail is a function of their distance from the surfactant head. For example, the CH_2 group closest to the head (group 13) has an f value of 0.41, and the next CH_2 group (group 14) has an f value of 0.25. The average f value of the groups in the OG tail is 0.24.

Although most of the hydrating contacts experienced by the atoms in the OG tail are made with water, these atoms also make a significant number of contacts with hydrogen-bonding atoms in the OG head (groups 1, 3, 5, 7, 10, and 12). The average value of f_{water} (−■−) of the groups in the OG tail is 0.18, and the average value of f_{head} (−▲−) of the groups in the OG tail is 0.07. Clearly, the value of $g_{tr,CS-MT}$ that is computed in the CS−MT model will depend strongly on whether or not hydrogen-bonding atoms in the surfactant head are modeled as contributing to hydration or, in other words, on whether f or f_{water} is used in eqs 2 and 4.

3.2. Dodecyl Maltoside (DM). The fractional degree of hydration of DM is plotted as a function of group number in Figure 4. Three different fractional hydration profiles (f, f_{water}, and f_{head}) are shown in the figure. The same general trends in f (see the −♦− results), f_{water} (−■−), and f_{head} (−▲−) that are observed for OG are observed for DM. It is interesting to note, however, that the average f value of the groups in the DM tail

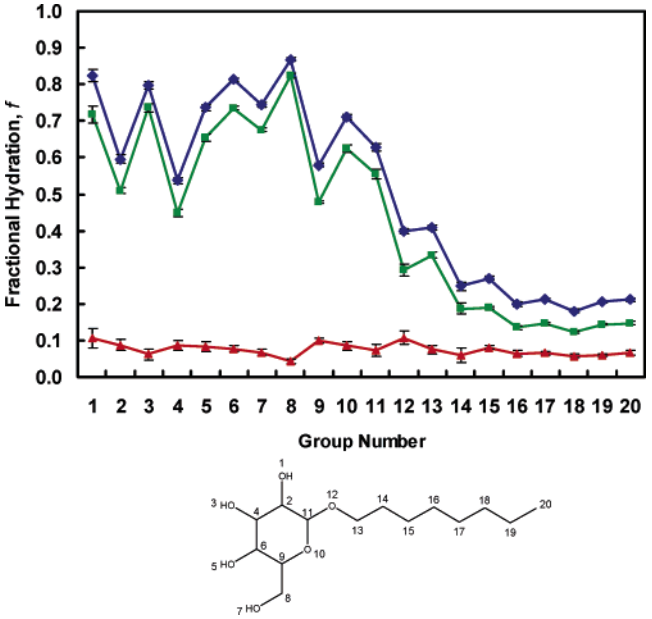


Figure 3. The average fractional degree of hydration, f, as defined in eq 3, for each of the groups in octyl glucoside (OG). Results are reported for fractional hydration values computed on the basis of counting contacts with water and with hydrogen-bonding groups in the surfactant head in the aggregate state (f, −◆−), for fractional hydration values on the basis of counting contacts only with water in the aggregate state (f_{water}, −■−), and for fractional hydration values on the basis of counting contacts only with hydrogen-bonding groups in the surfactant head in the aggregate state (f_{head}, −▲−). The chemical structure associated with each group is identified in the schematic of the molecule shown below the fractional hydration plot. The error bars shown correspond to the standard error of the mean.

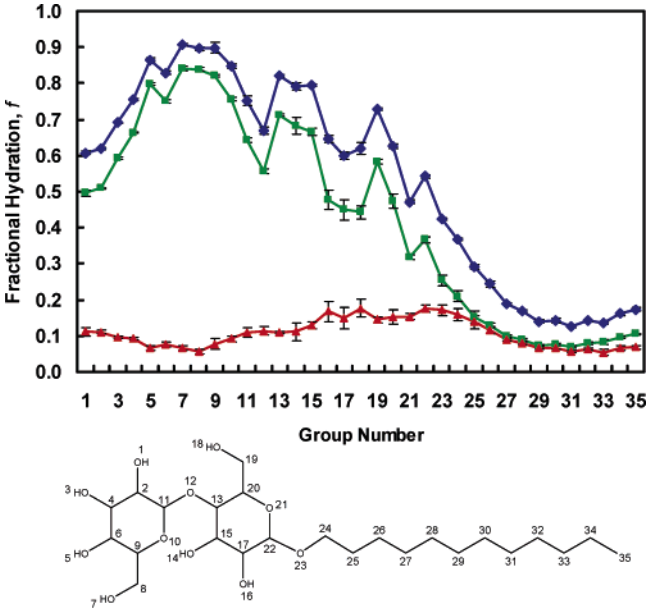


Figure 4. The average fractional degree of hydration, f, as defined in eq 3, for each of the groups in dodecyl maltoside (DM). Results are reported for fractional hydration values computed on the basis of counting contacts with water and with hydrogen-bonding groups in the surfactant head in the aggregate state (f, −◆−), for fractional hydration values on the basis of counting contacts only with water in the aggregate state (f_{water}, −■−), and for fractional hydration values on the basis of counting contacts only with hydrogen-bonding groups in the surfactant head in the aggregate state (f_{head}, −▲−). The chemical structure associated with each group is identified in the schematic of the molecule shown below the fractional hydration plot. The error bars shown correspond to the standard error of the mean.

(groups 24−35) is 0.19, which is significantly lower than the average f value of the groups in the OG tail (0.24). The DM tail is most likely less hydrated than the OG tail, on average, because the simulated DM micelle has a larger l_c value than the OG micelle, imparting to the hydrophobic core a lower surface area to volume ratio. The average f value of the groups in the DM head (groups 1−20) is 0.71, which is very similar to that of the OG head (0.69).

An interesting difference between the f_{head} profiles for OG and DM is that the average value of f_{head} for groups in the second ring structure in the DM head (groups 13−22), at 0.16, is significantly larger than the average value of f_{head} in the first ring of the DM head (groups 1−11), at 0.09. It is also significantly larger than the average value of f_{head} for the ring structure of the OG head (groups 1−11), at 0.08. The larger values of f_{head} experienced by groups 13−22 is most likely due to these groups being exposed to a higher concentration of hydrogen-bonding head groups than groups 1−11 in OG or in DM. Inspection of the contacts data shows that, although head contacts account for only 28% of the total hydrating contacts in OG, they account for 35% of the total hydrating contacts in DM. From these observations, the use of f_{water} values, rather than f values, in eqs 2 and 4 for CS−MT modeling is expected to have an even greater effect on the modeling results for DM than for OG.

3.3. Octyl Sulfinyl Ethanol (OSE). Fractional degree of hydration results (f, f_{water}, and f_{head}) for OSE are plotted as a function of group number in Figure 5. The average value of f (see the −◆− results) for the OSE head (0.73) is similar to that for the OG and the DM heads (0.69 and 0.71, respectively). However, the average value of f_{head} (−▲−) for OSE (0.04) is smaller than that for OG (0.08) and DM (0.11). This could be due to a lower concentration of hydrogen-bonding groups at the OSE micelle surface and to a lower affinity of the hydrogen-

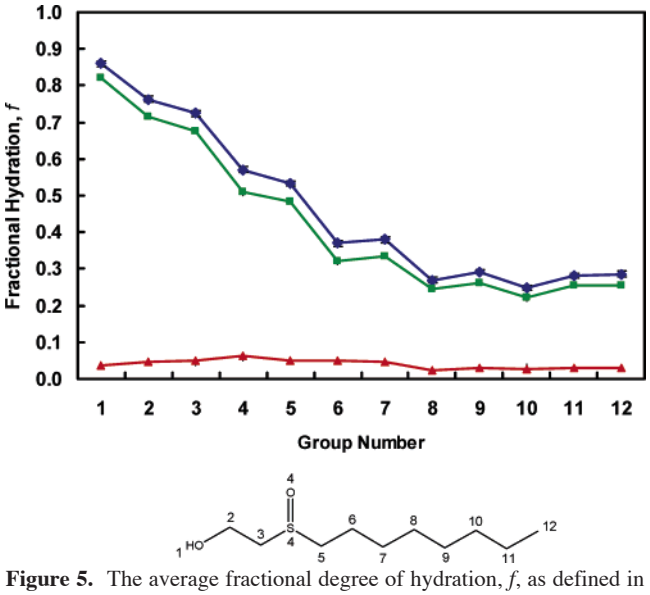


Figure 5. The average fractional degree of hydration, f, as defined in eq 3, for each of the groups in octyl sulfinyl ethanol (OSE). Results are reported for fractional hydration values computed on the basis of counting contacts with water and with hydrogen-bonding groups in the surfactant head in the aggregate state (f, −◆−), for fractional hydration values on the basis of counting contacts only with water in the aggregate state (f_{water}, −■−), and for fractional hydration values on the basis of counting contacts only with hydrogen-bonding groups in the surfactant head in the aggregate state (f_{head}, −▲−). The chemical structure associated with each group is identified in the schematic of the molecule shown below the fractional hydration plot. The error bars shown correspond to the standard error of the mean.

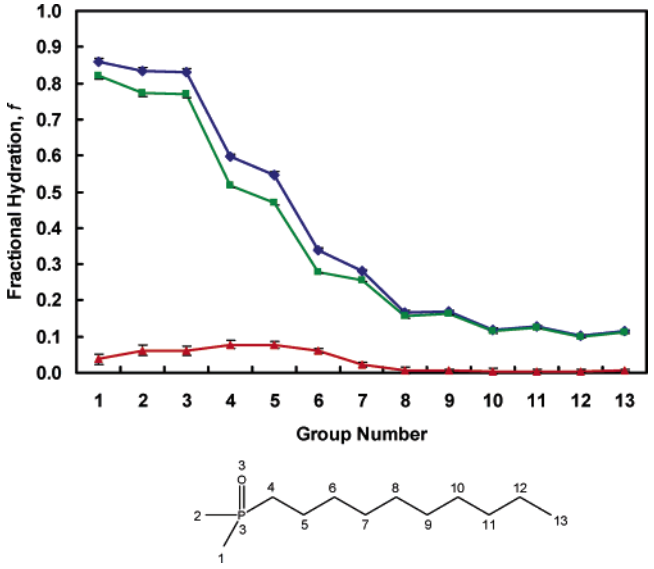


Figure 6. The average fractional degree of hydration, f, as defined in eq 3, for each of the groups in decyl dimethyl phosphine oxide ($C_{10}PO$). Results are reported for fractional hydration values computed on the basis of counting contacts with water and with hydrogen-bonding groups in the surfactant head in the aggregate state (f, −◆−), for fractional hydration values on the basis of counting contacts only with water in the aggregate state (f_{water}, −■−), and for fractional hydration values on the basis of counting contacts only with hydrogen-bonding groups in the surfactant head in the aggregate state (f_{head}, −▲−). The chemical structure associated with each group is identified in the schematic of the molecule shown below the fractional hydration plot. The error bars shown correspond to the standard error of the mean.

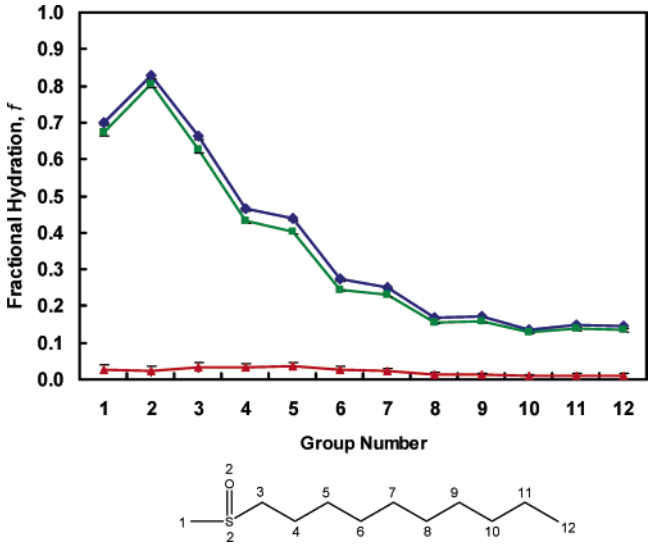


Figure 7. The average fractional degree of hydration, f, as defined in eq 3, for each of the groups in decyl ethyl sulfoxide ($C_{10}SO$). Results are reported for fractional hydration values computed on the basis of counting contacts with water and with hydrogen-bonding groups in the surfactant head in the aggregate state (f, −◆−), for fractional hydration values on the basis of counting contacts only with water in the aggregate state (f_{water}, −■−), and for fractional hydration values on the basis of counting contacts only with hydrogen-bonding groups in the surfactant head in the aggregate state (f_{head}, −▲−). The chemical structure associated with each group is identified in the schematic of the molecule shown below the fractional hydration plot. The error bars shown correspond to the standard error of the mean.

bonding groups for each other. The average value of f for the OSE tail groups (0.33) is more similar to that of OG (0.24) than to that of DM (0.19). It is interesting to note that the average value of f for the OSE tail groups is higher than that for the

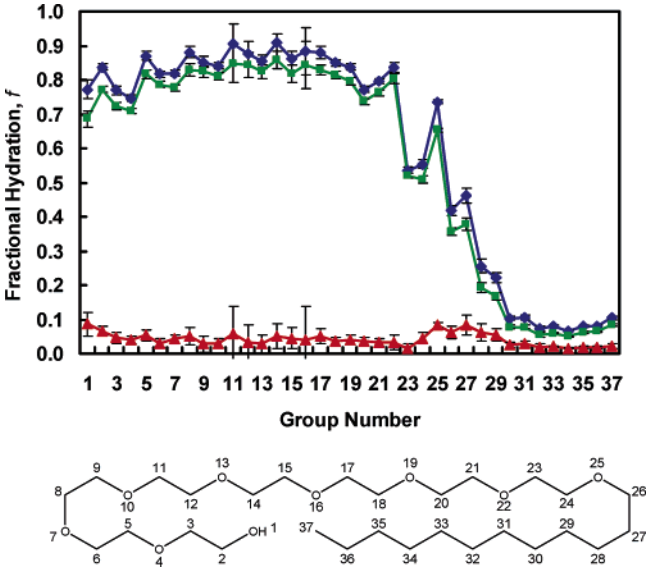


Figure 8. The average fractional degree of hydration, f, as defined in eq 3, for each of the groups in dodecyl octa(ethylene oxide) ($C_{12}E_8$). Results are reported for fractional hydration values computed on the basis of counting contacts with water and with hydrogen-bonding groups in the surfactant head in the aggregate state (f, −◆−), for fractional hydration values on the basis of counting contacts only with water in the aggregate state (f_{water}, −■−), and for fractional hydration values on the basis of counting contacts only with hydrogen-bonding groups in the surfactant head in the aggregate state (f_{head}, −▲−). The chemical structure associated with each group is identified in the schematic of the molecule shown below the fractional hydration plot. The error bars shown correspond to the standard error of the mean.

OG tail groups, indicating that the OSE micelle core is more hydrated than the OG micelle core, despite the fact that the simulated OSE micelle has a slightly smaller micelle core (l_c = 1.13 nm) than the simulated OG micelle (l_c = 1.19 nm). This comparatively high degree of hydration is due to the relatively large f values observed for hydrophobic groups 5, 6, and 7 in OSE relative to groups 13, 14, and 15 in OG.

It is interesting to note that the f values of groups 2 and 3, which are hydrophobic, are larger than that of group 4, which is hydrophilic. Clearly, the position of a group within a surfactant molecule (and therefore, relative to the micelle core), in addition to its chemical identity, is of importance in determining the degree of dehydration that it experiences upon micelle formation.

3.4. Decyl Dimethyl Phosphine Oxide ($C_{10}PO$). Fractional degree of hydration results (f, f_{water}, and f_{head}) for $C_{10}PO$ are plotted as a function of group number in Figure 6. The average f values of the $C_{10}PO$ head groups (groups 1−3) and the $C_{10}PO$ tail groups (groups 4−13) are 0.84 and 0.26, respectively (see the −◆− results). The degree of hydration of the $C_{10}PO$ head is higher than that observed in OG, DM, and OSE. It is interesting to note that groups 1, 2, and 3 have similar f values despite the hydrophobic character of groups 1 and 2. In addition, the average f value of the $C_{10}PO$ tail groups (0.26) is significantly larger than that of the DM tail groups (0.19), despite the fact that the simulated $C_{10}PO$ micelle is somewhat smaller (l_c = 1.52) than the simulated DM micelle (l_c = 1.56). This comparatively high degree of hydration is due to the relatively large f values observed for hydrophobic groups 4, 5, and 6 in $C_{10}PO$ relative to groups 24, 25, and 26 in DM.

3.5. Decyl Methyl Sulfoxide ($C_{10}SO$). Fractional degree of hydration results (f, f_{water}, and f_{head}) for $C_{10}SO$ are plotted as a function of group number in Figure 7. The average f values (see the −◆− results) of the $C_{10}SO$ head groups (groups 1 and 2) and the $C_{10}SO$ tail groups (groups 3−12) are 0.76 and 0.29,

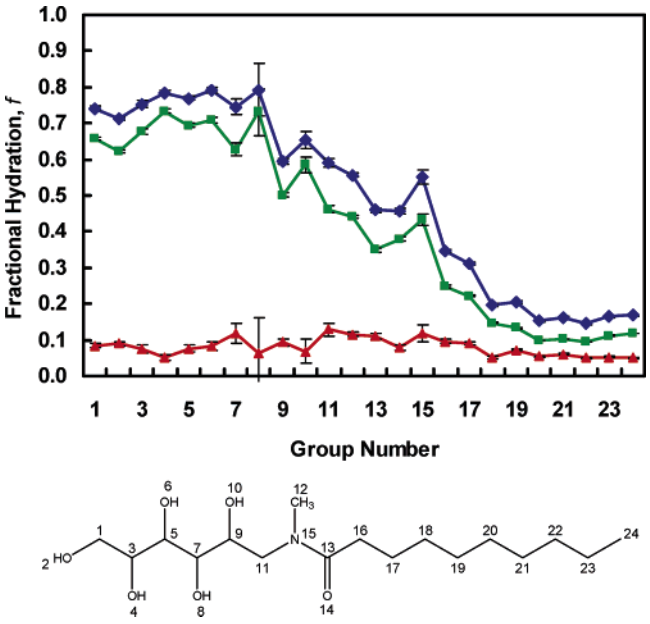


Figure 9. The average fractional degree of hydration, f, as defined in eq 3, for each of the groups in decanoyl-*n*-methylglucamide (MEGA-10). Results are reported for fractional hydration values computed on the basis of counting contacts with water and with hydrogen-bonding groups in the surfactant head in the aggregate state (f, −◆−), for fractional hydration values on the basis of counting contacts only with water in the aggregate state (f_{water}, −■−), and for fractional hydration values on the basis of counting contacts only with hydrogen-bonding groups in the surfactant head in the aggregate state (f_{head}, −▲−). The chemical structure associated with each group is identified in the schematic of the molecule shown below the fractional hydration plot. The error bars shown correspond to the standard error of the mean.

respectively. In contrast to the results obtained for CH_3 groups 1 and 2 in $C_{10}PO$, the hydrophobic CH_3 group 1 in $C_{10}SO$ has a lower f value than the hydrophilic groups in the head (groups 3 and 2 for $C_{10}PO$ and $C_{10}SO$, respectively). This difference between $C_{10}PO$ and $C_{10}SO$ may reflect differences in the atomic charges of the two surfactant heads. Despite the superficial chemical similarity of these two head groups, the atomic charges predicted using the CHelpG algorithm for the two heads are quite different, with $C_{10}PO$ having a charge distribution that makes its dipole moment roughly twice as large as that of $C_{10}SO$. Differences in the hydration profiles of the $C_{10}PO$ and the $C_{10}SO$ tails could be due both to differences in the head atomic charges and to differences in the shape and size of the two surfactant heads.

3.6. Dodecyl Octa(Ethylene Oxide) ($C_{12}E_8$). Fractional degree of hydration results (f, f_{water}, and f_{head}) for $C_{12}E_8$ are plotted as a function of group number in Figure 8. The average f values (see the −◆− results) of the $C_{12}E_8$ head groups (groups 1−25) and the $C_{12}E_8$ tail groups (groups 26−37) are 0.81 and 0.17, respectively. The small average degree of hydration of the hydrophobic core can be explained by the relatively large size of the micelle hydrophobic core (l_c = 1.51 nm), which is similar to that of the simulated DM micelle (l_c = 1.56 nm).

Because $C_{12}E_8$ and DM both have relatively large heads and the same hydrocarbon tail length, it is instructive to compare the fractional hydration results for $C_{12}E_8$ and DM. The most striking difference between the fractional hydration profiles of $C_{12}E_8$ and DM is the relatively low f_{head} values (−▲−) observed for the $C_{12}E_8$ head groups. The average value of f_{head} for the $C_{12}E_8$ head groups (groups 1−25), which have an average value of 0.04, is significantly smaller than the average value of f_{head} for the DM head groups (0.12). A possible explanation for this difference is that the hydrogen-bonding groups in the $C_{12}E_8$ head are not as attracted to each other as are the head groups in DM. The DM head contains both hydrogen bond donors and acceptors, and the $C_{12}E_8$ head contains several hydrogen bond acceptors but only a single hydrogen bond donor (OH group 1). Therefore, a DM head is capable of forming hydrogen bonds with both water and with other DM heads; a $C_{12}E_8$ head primarily forms hydrogen bonds with water, a difference that is expected to lower the average value of f_{head} of $C_{12}E_8$ relative to that of DM.

3.7. Decanoyl-*n*-Methylglucamide (MEGA-10). Fractional degree of hydration results (f, f_{water}, and f_{head}) for MEGA-10 are plotted as a function of group number in Figure 9. The average f value (see the −◆− results) of the groups in the MEGA-10 tail (groups 16−24) is 0.21, which is somewhat smaller than those in OG (0.24) and OSE (0.33). The average f value of the remaining groups in MEGA-10 (1−15) is 0.66. In general, the values of f_{head} (−▲−) observed for the MEGA-10 head groups are more similar to those of OG and DM than to that of $C_{12}E_8$. This can be understood by noting that the head structure of MEGA-10, like those of OG and DM, contains both hydrogen bond donors and acceptors.

MEGA-10 is difficult to model using the traditional MT modeling approach because it has a hydrophilic nitrogen atom (group 15) surrounded by three hydrophobic groups (CH_2 group 11, CH_3 group 12, and carbonyl groups 13 and 14). A logical starting point for traditional MT modeling would be to identify groups 16−24 as the MEGA-10 tail and to model each of the remaining groups as being part of the MEGA-10 head. By so doing, of course, the approximation is made that all the MEGA-10 head groups (including groups 11−14) remain fully hydrated in the micellar state. The hydration results presented in Figure 9 clearly show that this is indeed an approximation and that, because of their hydrophobic nature and location within the molecule, groups 11−13 are significantly more dehydrated than other hydrophilic or hydrophobic groups in the MEGA-10 head. In Section 4.9, we will compare the CMC predicted by the CS−MT model with the CMC predicted by the traditional MT model and test the assumption that groups 16−24 are part of the MEGA-10 tail.

3.8. Evaluation of Approximations Made in Traditional Molecular-Thermodynamic Modeling. The accuracy of two approximations made in traditional MT modeling to quantify the hydrophobic driving force for micelle formation can be evaluated using the computer simulation data presented above. These include (i) the accuracy of using surfactant head and tail assignments to compute g_{tr} and (ii) the extent to which surfactant heads shield the micelle core from hydrating contacts.

3.8.1. Accuracy of Computing g_{tr} using the Head and Tail Approximation. As discussed in detail in article 1,[1] in traditional MT modeling, the hydrophobic contribution to micelle formation is computed as the sum of two free-energy contributions: the transfer free-energy contribution, g_{tr}, and the interfacial free-energy contribution, g_{int}. A key assumption underlying the traditional MT modeling approach to compute g_{tr} is that the presence of the surfactant head at one end of the surfactant tail does not affect the change in hydration experienced by each group in the surfactant tail as it is transferred from the aqueous solution to a bulk solution of surfactant tails. By making this approximation, g_{tr} may be computed as if the surfactant heads are not present.

However, in practice, changes in the hydration state of the surfactant tail that occur upon its transfer to a bulk phase of tails are affected by the nature of the surfactant head attached

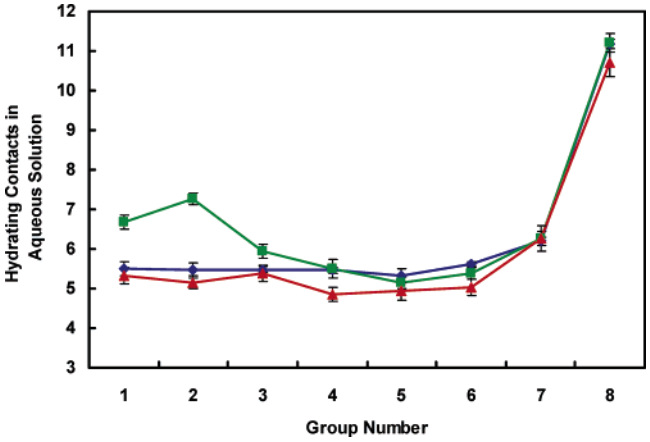


Figure 10. The number of hydrating contacts, as defined in the text, experienced in bulk water for 8 hydrophobic CH_2 or CH_3 groups in hexadecane (−◆−), octyl glucoside (OG, −■−), and octyl sulfinyl ethanol (OSE, −▲−). The first group listed for each surfactant corresponds to the CH_2 group adjacent to the tail's point of attachment to the surfactant head, and the last group listed for all three alkyl chains corresponds to the terminal CH_3 group. The error bars shown correspond to the standard error of the mean.

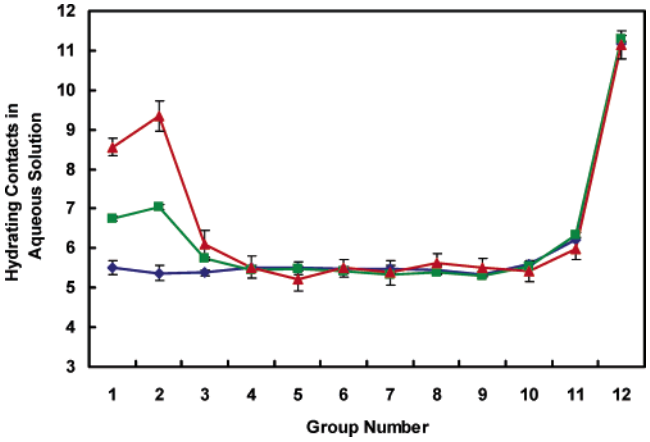


Figure 11. The number of hydrating contacts, as defined in the text, experienced in bulk water for 12 hydrophobic CH_2 or CH_3 groups in hexadecane (−◆−), dodecyl maltoside (DM, −■−), and dodecyl octa-(ethylene oxide) ($C_{12}E_8$, −▲−). The first group listed for each surfactant corresponds to the CH_2 group adjacent to the tail's point of attachment to the surfactant head, and the last group listed for all three alkyl chains corresponds to the terminal CH_3 group. The error bars shown correspond to the standard error of the mean.

to it. This can be shown by comparing hydration data for an oil molecule in bulk water to that of a surfactant tail in bulk water. We have computed the average number of hydrating contacts (as defined in Section 2.4) in bulk aqueous solution for each CH_2 and CH_3 group in hexadecane, as well as for those in the linear alkyl tails of several representative nonionic surfactants with tails containing either 8 or 12 CH_x ($x = 2$ or 3) groups. A cutoff of 0.3 nm was used to define hydrating contacts. In Figure 10, plots of hydrating contacts as a function of group number in the bulk aqueous solution are shown for 8 hexadecane groups (see the −◆− results), the 8 hydrophobic tail groups of OG (−■−), and the 8 hydrophobic tail groups of OSE (−▲−). In Figure 10, group 1 for OG and OSE is defined as the CH_2 group adjacent to the surfactant head, and group 8 for OG, OSE, and hexadecane corresponds to the terminal CH_3 group of each alkyl chain. In Figure 11, plots of hydrating contacts as a function of group number in the bulk aqueous solution are shown for 12 hexadecane groups (−◆−), the 12 hydrophobic tail groups of DM (−■−), and the 12 hydrophobic tail groups of $C_{12}E_8$ (−▲−). In Figure 11, group 1 for the surfactants is defined as the CH_2 group adjacent to the surfactant head, and group 12 for the surfactants and for hexadecane corresponds to the terminal CH_3 group of each alkyl chain. Note that we have not compared the OG and the OSE results with octane hydration data, or the DM and the $C_{12}E_8$ results with dodecane hydration data, because group 1 for each surfactant is a CH_2 group, rather than a CH_3 group. The hydration data presented in Figures 10 and 11 for the surfactants and hexadecane enables us to compare the hydration of chemically identical hydrophobic groups. The error bars shown in Figures 10 and 11 represent standard errors of the mean.

A comparison of the hydration profiles for OG, OSE, DM, and $C_{12}E_8$ with corresponding groups in hexadecane clearly shows that the surfactant head type has an effect on the hydration state of the adjacent CH_2 group in aqueous solution. In addition, our results show that the presence of the surfactant head affects the hydration states of a significant number of hydrophobic groups further down each linear alkyl chain (groups 2 and 3 for OG, DM, and $C_{12}E_8$; groups 2−6 for OSE). The total number of hydrating contacts experienced by groups 1−8 of OG and OSE is 6.3% larger and 5.16% smaller, respectively, than those experienced by the 8 corresponding groups in hexadecane. The total number of hydrating contacts experienced by groups 1−12 of DM and $C_{12}E_8$ is 4.2% larger and 10.1% greater, respectively, than those experienced by the 12 corresponding groups in hexadecane.

Because the extent of hydration of the surfactant tail in the bulk water reference state is affected by the type of surfactant head, it follows that the change in hydration incurred upon transfer of the surfactant tail to a bulk phase of tails is affected by the type of surfactant head. The assumption made in traditional MT modeling in computing g_{tr} that every tail group is dehydrated to the same degree as a tail with no attached head is clearly an approximation.[2,8] Nevertheless, it is important to point out that traditional MT modeling has been shown to yield quantitatively, or semiquantitatively, accurate predictions of the micellization behavior of relatively simple nonionic surfactants.[2,8] Accordingly, the hydration approximations that are made in traditional MT modeling to compute g_{tr} appear reasonable in the absence of detailed hydration data. The CS−MT modeling approach, however, eliminates the need to make such approximations by computing the changes in hydration that occur upon micelle formation directly from molecular dynamics simulation results.

3.8.2. Effect of the Surfactant Heads on Aggregate Core Hydration. The interfacial free-energy contribution, g_{int}, reflects the free-energy penalty associated with forming the micelle core−water interface and is computed in the context of traditional MT modeling using a micelle core−water interfacial tension. Specifically, g_{int} is computed as follows:[1]

$$g_{int} = (a - a_0)\sigma \tag{11}$$

where a is the area per surfactant molecule at the micelle core−water interface, a_0 is the interfacial area that is screened by each surfactant head, and σ is a composition-weighted average of the curvature-corrected interfacial tension between water and a bulk phase of hydrophobic tails (for complete details, see article 1). In our past work, we have estimated a_0 as being equal to 21 Å^2 for every surfactant head in the micelle,[8] an area that is equal to the cross-sectional area of a linear alkyl chain. Nagarajan et al. have modeled a_0 as being equal to the smaller of two areas−

TABLE 2: Evaluation of Traditional MT Modeling Approximations[a]

	N_{cont}	$N_{cont} - N_{cont,oil\ agg}$	shielded area [Å^2]
octane	18.25	0.00	
OG	10.65	−7.61	31.22
OSE	15.28	−2.97	12.19
dodecane	18.97	0.00	
DM	12.14	−6.83	26.99
$C_{12}E_8$	14.66	−4.31	17.03

[a] Simulation results for the number of hydrating contacts (N_{cont}) on a per oil or surfactant molecule basis experienced by octane, octyl glucoside (OG), octyl sulfinyl ethanol (OSE), dodecane, dodecyl maltoside (DM), and dodecyl octa(ethylene oxide) ($C_{12}E_8$) in the aggregate environment. The octane results correspond to the results for a spherical octane aggregate of aggregation number 25.[1] The dodecane results correspond to a spherical dodecane aggregate of aggregation number 33.[1] As described in the text, hydrating contacts results for each micelle have been scaled by the ratio of the micelle core area to the area of either the octane or the dodecane oil aggregate. The change in the number of hydrating contacts relative to the corresponding oil aggregate ($N_{cont} - N_{cont,oil\ agg}$) is also listed, again on a per oil or per surfactant molecule basis. The shielded area reported for the surfactant micelles corresponds to the area at the micelle core−water interface that is effectively shielded from hydrating contacts by the surfactant heads.

the cross-sectional area of a linear alkyl chain (21 Å^2) or the cross-sectional area of the surfactant head.

The curvature-dependent interfacial tension of a surfactant tail of type j in the micelle has been estimated using the Gibbs−Tolman−Koenig−Buff equation:[28−31]

$$\sigma_j = \frac{\sigma_{0,j}}{\left(1 + \frac{(S-1)\delta}{l_c}\right)} \tag{12}$$

where $\sigma_{0,j}$ is the interfacial tension of component j at a flat interface with water (typically around 50 mN/m for hydrocarbons), δ is the Tolman distance, and S is a shape factor that is equal to 3 for spheres, 2 for cylinders, and 1 for disks or bilayers. The estimation of $\sigma_{0,j}$ for alkyl chains of varying length and as a function of temperature, as well as the estimation of the Tolman distance, δ, were discussed in detail in article 1.[1]

The simulation data reported in Sections 3.1−3.7 can be used to evaluate how physically reasonable eq 11 is as implemented in the context of the traditional MT modeling approach. By comparing simulation results for the small spherical octane aggregate (see article 1)[1] with those for the OG and the OSE micelles, as well as simulation results for the large spherical dodecane aggregate[1] with those for the DM and the $C_{12}E_8$ micelles, it is possible to comment on the extent to which the surfactant heads shield the micelle hydrophobic core from hydrating contacts. These two oil aggregates have been selected for comparison because their surface areas are similar to those of the simulated micelles. Table 2 reports the total number of hydrating contacts, N_{cont}, for octane, OG, OSE, dodecane, DM, and $C_{12}E_8$ in the aggregate state on a per molecule basis. To allow direct comparison of the results for the oil aggregates and for the micelles, N_{cont} for OG and OSE were computed by scaling the OG and the OSE hydrating contact results by ($A_{core,mic}$/$A_{core,oct\ agg}$), where $A_{core,mic}$ is the surface area of the surfactant micelle core and $A_{core,oct\ agg}$ is the surface area of the octane aggregate, to correct for differences in aggregate surface area. We consider scaling based on surface area to be appropriate because hydrating contacts should be approximately proportional to the exposed surface area. Similarly, N_{cont} for DM and $C_{12}E_8$ was computed by scaling the DM and the $C_{12}E_8$ hydrating contact results by ($A_{core,mic}$/$A_{core,dod\ agg}$), where $A_{core,dod\ agg}$ is the surface area of the dodecane aggregate. Table 2 also reports the difference between the number of hydrating contacts for OG or OSE and the number of hydrating contacts for octane, as well as the difference between the number of hydrating contacts for DM or $C_{12}E_8$ and the number of hydrating contacts for dodecane as ($N_{cont} - N_{cont,oil\ agg}$). The decrease in the number of hydrating contacts observed for each surfactant micelle relative to the corresponding value for each oil aggregate can be used to infer the area at the micelle core−water interface that is effectively shielded from hydrating contacts by the surfactant heads. These results are reported in Table 2 as the "shielded area" values. The average of these shielded area values (21.85 Å^2) is very similar to the shielded area value used in traditional MT modeling (21 Å^2). However, inspection of the shielded area results reveals that the shielded area varies significantly among the surfactants considered. The shielded area calculated for OSE, for example, at 12.19 Å^2, is less than half of the shielded area calculated for OG (31.22 Å^2). An advantage of the CS−MT modeling approach is that it permits estimation of the hydrophobic contribution to g_{form} (as reflected in $g_{dehydr} + g_{hydr}$) without making any assumptions about the effect of the surfactant heads on the hydration state of the micelle core.

4. Molecular-Thermodynamic Modeling Based on Computer Simulation Inputs

4.1. Using the CS−MT Modeling Approach to Predict Surfactant Micellization Behavior. As stressed in article 1, using the CS−MT modeling approach to quantify the hydrophobic effect for oil aggregates in water is less challenging than using it to model surfactant micellization. Indeed, in surfactant micellization, the presence of the surfactant heads at the aggregate core−water interface introduces several complications that are absent in the oil aggregate case. To deal with these complications, we will make a number of simplifying approximations. Following a description of these approximations in Sections 4.1.1 and 4.1.2, we will use the CS−MT model to predict g_{form} and the CMC for each of the seven nonionic surfactants discussed in Section 3. Although the CS−MT model enables the prediction of a variety of micellar solution properties from g_{form} (including micelle shape, size, and composition), the surfactant CMC was selected for prediction and comparison with experimental CMC data because the CMC depends exponentially on g_{form}, and as such, it provides a stringent quantitative test with which to evaluate the predictive accuracy of the CS−MT model. The CMCs predicted by the CS−MT model will be compared both with the CMCs predicted by the traditional MT model and with the experimental CMC values. In addition, the accuracy of the approximations made in implementing the CS−MT model will be discussed in the context of the CS−MT and the traditional MT modeling results.

4.1.1. Estimation of g_{dehydr}. When using the CS−MT model to quantify the hydrophobic effect, eq 2 is used for every hydrophobic group in the surfactant molecule, regardless of whether the hydrophobic group is part of the surfactant head or the surfactant tail. Accordingly, every hydrophobic group in the surfactant molecule contributes to the hydrophobic driving force for micelle formation to the extent that the group is dehydrated upon micelle formation. To implement eq 2, therefore, suitable values of g_{tr_i} must be estimated for every hydrophobic group in the surfactant molecule.

In traditional MT modeling, only the surfactant tails are considered to be dehydrated upon micelle formation. In that case, the transfer free-energy contribution, g_{tr}, of these tails can

be estimated in a straightforward manner using experimental tail solubility data, or a theoretical estimate of tail solubility can be made using a group-contribution approach (see article 1 for a detailed discussion).[1] Similarly, in CS−MT modeling, g_{tr_i} values of hydrophobic groups in the surfactant tail can be determined from an estimate of the solubility of group *i* in water. Suitable g_{tr_i} values of hydrophobic groups in the surfactant head are more difficult to estimate. Dehydration of the surfactant heads upon micelle formation may result from solvent exclusion by other surfactant heads, contact with hydrophobic groups in other surfactant heads, and contact with the micelle hydrophobic core. Due to the highly anisotropic nature of the micelle core−water interfacial region, it is difficult to assign suitable g_{tr_i} values associated with transfer from bulk water to this interfacial environment because the required experimental solubility data is not available. To the best of our knowledge, there are no simple theoretical approaches to predict g_{tr_i} between bulk water and such a complex, anisotropic environment. A second complication in estimating g_{tr_i} values for groups in the surfactant head results from the fact that the hydrophobicity of hydrophobic groups in the surfactant head may be affected by their being bonded to hydrophilic groups. In molecular mechanics force fields, the chemical effect of being bonded to a hydrophilic group is captured by the fact that the atomic charge of each atom in a molecule is a function of its adjacent atoms.[11] For example, in the OPLS-AA force field, the net atomic charge assigned to the CH_2 group in poly(ethylene oxide) is 0.1 | *e* | (where *e* is the charge of an electron), which differs from the net charge of 0 | *e* | assigned to a CH_2 group bonded to alkyl groups.

With the above complications in mind, in order to implement the CS−MT model in a straightforward manner, in this article, we make the approximation that the g_{tr_i} values of hydrophobic groups in the surfactant head are identical to the g_{tr_i} values of the same hydrophobic groups in the surfactant tail. We anticipate that implementing the CS−MT model in this approximate way should yield an improvement over the traditional MT modeling approach for many surfactants. In the traditional MT modeling approach, hydrophobic groups in the surfactant head do not contribute at all to the hydrophobic driving force for micelle formation. However, it is important to note that obtaining better estimates of g_{tr_i} for hydrophobic groups in the surfactant head represents an important area for future research in order to improve the accuracy of the CS−MT model.

For the seven nonionic surfactants modeled here, we only need to consider CH, CH_2, and CH_3 hydrophobic groups. Accordingly, to implement the CS−MT model, g_{tr_i} values for CH_2 and CH_3 were estimated using the same solubility correlations for linear alkyl tails that are used in traditional MT modeling.[32] The g_{tr_i} value for CH was estimated using solubility data for branched alkyl tails.[33]

4.1.2. Estimation of g_{hydr}. In article 1, we presented an approach to theoretically estimate g_{hydr} for oil molecules. In our model for g_{hydr}, we calculated Δg_{wc} using eqs 5 and 6. In estimating g_{hydr} for each of the hydrophobic groups in a surfactant molecule, two complications arise: (i) g_{hydr} is nonzero only for those hydrophobic groups in the surfactant molecule that are adsorbed onto, or that penetrate into, the micelle hydrophobic core, and (ii) for those hydrophobic groups that are adsorbed onto, or that penetrate into, the micelle hydrophobic core, Δg_{wc} may be affected by the presence of the surfactant heads at the micelle core−water interface.

Note that the free-energy contribution, g_{hydr}, is zero for those hydrophobic groups that are not part of the micelle hydrophobic core, because g_{hydr} accounts for the difference in free energy associated with hydrating contacts in the bulk water and in the aggregate environment. Fundamentally, the origin of g_{hydr} is the size dependence of hydration thermodynamics. An isolated hydrophobic chain in water is much smaller in size than a typical aggregate core−water interface. Therefore, the hydrophobic chain disrupts the hydrogen-bonding and coordinate-bonding network of the aqueous solution to a different extent. Hydrophobic groups in a surfactant molecule that are not part of the micelle hydrophobic core continue to disrupt this hydrogen-bonding and coordinate-bonding network in the aggregate state in much the same way that they do in the bulk water state. Consequently, g_{hydr} for such groups is zero. With this in mind, we will consider any hydrophobic group that has an *f* value equal to, or less than, 0.60 to be part of the micelle hydrophobic core (see below) and to have a nonzero value of g_{hydr}.

The selection of $f = 0.60$ as the appropriate cutoff value was motivated both by physical intuition and by our computer simulation results. For a hydrophobic group located precisely at a flat oil−water interface (with half of its surface in oil and the other half in water), the average value of *f* computed through molecular dynamics simulation would be 0.5. For a hydrophobic group adsorbed at a curved, rough oil−water interface, we would expect the average value of *f* to be greater than 0.5. We note that, for the small, spherical hexadecane oil aggregate simulated in article 1, the average value of *f* for the two terminal CH_3 groups was found to be 0.53, even though each of the CH_3 and the CH_2 groups in the hexadecane molecules are part of the aggregate core. We selected $f = 0.6$ as a suitable cutoff value after considering simulation results for the seven nonionic surfactants modeled here, as well as simulation results for a number of simple, ionic surfactants, including sodium dodecyl sulfate (SDS), cetyltrimethylammonium bromide (CTAB), and decyltrimethylammonium bromide (DTAB). CS−MT modeling results for these three ionic surfactants will be presented in article 3 of this series.[34] For each of the nonionic and ionic surfactants that we simulated, groups that would be considered to be part of the surfactant head in traditional MT modeling had an *f* value greater than 0.60, and groups that would be considered to be part of the surfactant tail in traditional MT modeling had an *f* value that is less than, or equal to, 0.60. In traditional MT modeling, all the atoms in the surfactant tail are considered to be part of the micelle core, and such an assignment yields quantitatively, or semiquantitatively, accurate predictions of the micellar solution behavior of simple surfactants.[8] In this article, therefore, we treat any hydrophobic groups in a surfactant molecule with an *f* value that is less than, or equal to, 0.60 as being part of the micelle hydrophobic core.

For those hydrophobic groups that are identified as being part of the micelle hydrophobic core, a reasonable value of Δg_{wc} must be estimated in order to compute g_{hydr} using eq 4. We propose that, to a first approximation, Δg_{wc} can be evaluated as being equal to the value calculated for oil molecules in article 1.[1] As shown in eq 5, Δg_{wc} is the difference between the free energy per unit SASA in the aggregate core state (σ_{core}) and in the bulk water state (σ_{bulk}). The value of σ_{bulk} for a hydrophobic group in a surfactant molecule is very similar to that of σ_{bulk} for a hydrophobic group in an oil molecule. Furthermore, the success of the traditional MT approach in modeling the aggregate core−water interface using an oil−water interfacial tension[1] indicates that assuming that σ_{core} in a micelle is equal to σ_{core} of an oil−water interface is a reasonable approximation. This approximation has been made in modeling each of the

TABLE 3: Modeling Results for the Simulated Micelles[a]

surfactant	g_{dehydr} [k_BT]	g_{hydr} [k_BT]	$\hat{g}_{int}$ [k_BT]	$g_{tr,CS-MT}$ [k_BT]	g_{tr} [k_BT]
OG	-12.11 ± 0.06	1.51 ± 0.02	4.01	-14.61 ± 0.06	-13.98
DM	-17.90 ± 0.05	2.25 ± 0.02	4.71	-20.36 ± 0.06	-20.06
OSE	-10.53 ± 0.10	1.31 ± 0.02	4.34	-13.55 ± 0.10	-13.98
$C_{10}PO$	-14.34 ± 0.05	1.44 ± 0.03	3.86	-16.75 ± 0.06	-16.96
$C_{10}SO$	-13.08 ± 0.12	1.38 ± 0.04	3.86	-15.56 ± 0.12	-16.96
$C_{12}E_8$, all hydr groups	-21.04 ± 0.22	1.29 ± 0.06	4.90	-24.65 ± 0.23	-19.95
$C_{12}E_8$, tail hydr groups	-16.76 ± 0.09	0.55 ± 0.05	4.90	-21.10 ± 0.10	-19.95
MEGA−10	-14.91 ± 0.16	1.36 ± 0.06	3.81	-17.36 ± 0.17	-15.47

[a] CS−MT and traditional MT modeling results for each of the seven simulated nonionic surfactant micelles considered in this article. CS−MT model predictions of g_{dehydr}, g_{hydr}, $\hat{g}_{int}$, and $g_{tr,CS-MT}$ were made as described in Section 1.1. The uncertainties reported for the CS−MT model predictions correspond to the standard error of the mean. Traditional MT modeling results for g_{tr} are presented for comparison with $g_{tr,CS-MT}$. For $C_{12}E_8$, CS−MT modeling results generated by summing in eq 2 over all the hydrophobic (hydr) groups in the $C_{12}E_8$ molecule, as well as over only the $C_{12}E_8$ tail hydrophobic (hydr) groups, are reported separately (see Section 4.8).

TABLE 4: Modeling Results for the Optimal Micelles[a]

surfactant	shape	n	g_{int} [k_BT]	g_{pack} [k_BT]	g_{st} [k_BT]	g_{form} [k_BT] (CMC [mM]) CS−MT model	MT model	experimental
OG	cyl	43	3.23	2.17	1.54	-8.67 ± 0.06 (9.52 ± 0.57)	-8.04 (17.97)	-7.74 (24.1)
DM	cyl	58	4.31	2.23	1.91	-12.89 ± 0.06 (0.14 ± 0.01)	-12.60 (0.19)	-12.87 (0.14)
OSE	cyl	535	2.71	2.39	1.07	-8.39 ± 0.10 (12.59 ± 1.27)	-8.81 (8.26)	-7.62 (27)
$C_{10}PO$	cyl	45	4.08	2.16	1.87	-9.64 ± 0.06 (3.61 ± 0.23)	-9.85 (2.93)	-9.58 (3.8)
$C_{10}SO$	cyl	12802	2.94	2.50	1.0	-10.13 ± 0.12 (2.23 ± 0.28)	-11.53 (0.54)	-10.38 (1.7)
$C_{12}E_8$	cyl	54	4.41	2.23	1.95	all hydr groups -17.06 ± 0.23 (0.002 ± 0.0) tail hydr groups -13.51 ± 0.10 (0.08 ± 0.01)	-12.36 (0.24)	-13.22 (0.1)
MEGA−10	sph	22	4.96	2.39	1.97	-9.05 ± 0.17 (6.55 ± 1.15)	-8.12 (43.33)	-9.31 ± 0.01 (5 ± 0.05)

[a] CS−MT and traditional MT modeling results for each of the seven nonionic surfactants considered in this article. Both the CS−MT model and the traditional MT model yield identical predictions for the optimal micelle shape, the number-average aggregation number (n), g_{int}, g_{pack}, and g_{st} (see Section 4.2). The CS−MT and the traditional MT model predictions of g_{form} were obtained using the values of $g_{tr,CS-MT}$ and g_{tr} reported in Table 3, respectively, as an input to eq 8. The CS−MT and the traditional MT model predictions of the CMC and the value of g_{form} inferred using the experimental CMC data were calculated using eq 9. The uncertainties reported for the CS−MT model predictions correspond to the standard error of the mean. For $C_{12}E_8$, CS−MT modeling results generated by summing in eq 2 over all the hydrophobic (hydr) groups in the $C_{12}E_8$ molecule, as well as over only the $C_{12}E_8$ tail hydrophobic (hydr) groups, are reported separately (see Section 4.8).

seven nonionic surfactants considered in this article, which we discuss below.

4.2. Modeling Results for Octyl Glucoside (OG). Using the simplifying approximations discussed in Sections 4.1.1 and 4.1.2, we used the CS−MT model to predict the micellization behavior for OG in aqueous solution at 25 °C. In Table 3, we report CS−MT modeling results for the *simulated* OG micelle, including (i) g_{dehydr}, (ii) g_{hydr}, (iii) $\hat{g}_{int}$, and (iv) $g_{tr,CS-MT}$. The reported uncertainty for the CS−MT modeling results corresponds to the standard error of the mean, as computed through block averaging. The CS−MT modeling results for g_{dehydr} and g_{hydr} were generated using f values, which, as discussed in Section 3.1, are based on contacts with water and with hydrogen-bonding groups in the surfactant head. The value of g_{dehydr} ($-12.11k_BT$) is much larger in magnitude than that of g_{hydr} ($1.51k_BT$). The physical interpretation of this result is that the free-energy contribution associated with the dehydration of each group in OG is significantly larger in magnitude than the free-energy contribution arising from the difference in free energy associated with hydration in the micellar state and in the bulk aqueous solution (or, in other words, the free-energy contribution arising from the size dependence of hydration thermodynamics discussed in Section 3.4 in article 1). However, as shown in Table 4 (see below), if g_{hydr} is not included in the CS−MT model, accurate CS−MT modeling results would not be obtained. Equation 7 was used to compute $g_{tr,CS-MT}$ from g_{dehydr}, g_{hydr}, and $\hat{g}_{int}$. In Table 3, we also report the traditional MT model prediction of g_{tr} for comparison with $g_{tr,CS-MT}$. We note that the CS−MT model prediction for the transfer free-energy contribution ($g_{tr,CS-MT} = -14.61k_BT$) is $0.63k_BT$ more negative than the traditional MT model prediction ($g_{tr} = -13.98k_BT$).

In Table 4, we report CS−MT and traditional MT modeling results for micelles of the optimal shape and size. The optimal micelle shape and size are predicted to be the values that minimize g_{form}.[8] Although the CS−MT model predicts a different g_{form} value than that obtained using the traditional MT model, both models yield identical predictions for the optimal micelle shape and size. This equivalence arises because the only contribution to g_{form} that differs in the two models (the transfer free-energy contribution) does not depend on the micelle shape and size and therefore does not affect the minimization procedure used to determine the optimal micelle properties. As discussed in Section 2.3, the computer simulation of the OG micelles was conducted for a spherical micelle with an aggregation number of 29, but both the CS−MT model and the traditional MT model predict that the optimal micelles are cylindrical with a number-average micelle aggregation number of 43. In Table 4, we report predictions of the CS−MT model and the traditional MT model for (i) the optimal micelle shape, (ii) the number-average aggregation number (n), (iii) g_{int}, (iv) g_{pack}, (iv) g_{st}, (v) the CS−MT model predictions of g_{form} and the CMC, (vi) the traditional MT model predictions of g_{form} and the CMC, and (vii) the experimental values of g_{form} and the CMC.[35] The reported uncertainty for the CS−MT modeling results is the standard error of the mean, as computed through block averaging. The CS−MT and the traditional MT model predictions for g_{form} were obtained by using $g_{tr,CS-MT}$ and g_{tr}, respectively, as an input to eq 8. In applying the CS−MT model

and the traditional MT model, the surfactant head area, a_h, was modeled as being equal to 40 Å^2.[8] Note that the molecular parameter, a_h, is used to calculate the steric free-energy contribution, g_{st}.[8,36] Traditional MT modeling results were generated using the traditional MT modeling approach reviewed in article 1.[1] In generating the traditional MT modeling results, each OG surfactant was modeled as having 7 CH_2 groups and 1 CH_3 group in the surfactant tail (groups 13−20 in Figure 3). The CS−MT and the traditional MT model predictions of the CMC and the value of g_{form} inferred using the experimental CMC data were calculated using eq 9.

Because the shape and size of the optimal micelles predicted by the CS−MT model and the traditional MT model are identical, the values of g_{int}, g_{pack}, and g_{st} predicted by each model are also identical.[8] Although the predicted value of g_{int} is the same in both the CS−MT model and the traditional MT model, as shown in Tables 3 and 4, the value of g_{int} computed for the optimal micelle ($3.23k_BT$) is significantly lower than the value of $\hat{g}_{int}$ computed for the simulated micelle ($4.01k_BT$). The free-energy contributions, g_{st} ($1.54k_BT$) and g_{pack} ($2.17k_BT$), although smaller than g_{int}, both contribute significantly to g_{form}. Values of g_{ent} and g_{elec} are not reported because they are equal to zero for this nonionic, single-surfactant system. The CS−MT model, the MT model, and the experimental values of g_{form} are all within $0.63k_BT$ of each other. The CS−MT and the traditional MT model predictions of the CMC, as well as the value of g_{form} inferred using the experimental CMC data, were calculated using eq 9. Both the CS−MT model and the traditional MT model predict CMC values that are somewhat lower than the experimental CMC value of 24.1 mM.[35] The CMC predicted by the CS−MT model is 61% lower than the experimental CMC value, and the CMC predicted by the traditional MT model is 25% lower than the experimental CMC value. This discrepancy reflects the different estimates of the hydrophobic driving force for micelle formation obtained using the two models ($g_{tr,CS-MT}$ and g_{tr}, as reported in Table 3). Although the traditional MT model result for the CMC is closer to the experimental CMC value than the CS−MT result for the CMC, we consider both the CS−MT model and the traditional MT model CMC predictions shown in Table 3 to be in reasonable agreement with the experimental data, given the exponential dependence of the CMC on g_{form} (see eq 9).

4.3. Modeling Results for Dodecyl Maltoside (DM). CS−MT modeling results for the simulated DM micelle are reported in Table 3, where each free-energy contribution was calculated as described in Section 4.2. Theoretical predictions for the optimal micelles obtained using the CS−MT model and the traditional MT model, as well as the experimental data[37] for the micellization behavior of DM in aqueous solution at 25 °C with 0.1 M of added NaCl, are reported in Table 4. The approach described in Section 4.2 was used to calculate each free-energy contribution, the g_{form} values, and the CMC values. In using the CS−MT model and the traditional MT model, the surfactant head area, a_h, was modeled as being equal to 52 Å^2.[8]

Although computer simulation of DM was conducted in a micelle with an aggregation number of 45, the optimal DM micelles that are predicted to form in solution by the CS−MT model and the traditional MT model are small cylinders with a number-average aggregation number of 58. The predicted value of g_{int} ($4.31k_BT$) is slightly lower than that of $\hat{g}_{int}$ ($4.71k_BT$) due to this aggregation number difference. The predicted value of g_{st} is slightly larger for the optimal DM micelle ($1.91k_BT$) than the predicted value for the optimal OG micelle ($1.54k_BT$), because the DM head (groups 1−23) is modeled as being 12 Å^2 larger in cross-sectional area than the OG head (groups 1−12).[8] As in OG, the CS−MT model prediction of the transfer free-energy contribution ($g_{tr,CS-MT} = -20.36k_BT$) is slightly more negative than that of the traditional MT model prediction ($g_{tr} = -20.06k_BT$). This leads to the CS−MT model predicting a lower CMC (0.14 mM) than that predicted by the traditional MT model (0.19 mM). In this case, the CS−MT model prediction of the CMC agrees remarkably well with the experimental value (0.14 mM).

4.4. Modeling Results for Octyl Sulfinyl Ethanol (OSE). CS−MT modeling results for the simulated OSE micelle are reported in Table 3, where each free-energy contribution was calculated as described in Section 4.2. Theoretical predictions for the optimal micelles obtained using the CS−MT model and the traditional MT model, as well as the experimental data[38] for the micellization behavior of OSE in aqueous solution at 25 °C, are reported in Table 4. The approach described in Section 4.2 was used to calculate each free-energy contribution, the g_{form} values, and the CMC values listed in Table 4. In applying the CS−MT model and the traditional MT model, the surfactant head area, a_h, was modeled as being equal to 30 Å^2.[8]

Although computer simulation of an OSE micelle was conducted at an aggregation number of 25, the optimal OSE micelles that are predicted using the CS−MT model are cylinders with a number-average aggregation number of 535. The predicted value of g_{int} ($2.71k_BT$) is significantly lower than that of $\hat{g}_{int}$ ($4.34k_BT$) due to this large aggregation number difference. The g_{int} value for OSE is $1.60k_BT$ lower than that for DM and $0.52k_BT$ lower than that for OG due to the large aggregation number of the OSE micelle, which lowers the interfacial area per surfactant molecule.[8] For OSE, $g_{tr,CS-MT}$ ($-13.55k_BT$) is slightly less negative than g_{tr} computed using the traditional MT modeling approach ($-13.98k_BT$). As a result, the CS−MT model prediction of the CMC is higher than the MT model prediction of the CMC, and it is also slightly closer to the experimental CMC value. As shown in Figure 5, groups 2 and 3 in OSE both have relatively high f values. However, the net effect of using eqs 2 and 4 to determine the contribution of both of these groups to the hydrophobic driving force for micelle formation is still significant, at $-0.71k_BT$. Allowing all the hydrophobic groups in the surfactant molecule (and not just those in the surfactant tail) to contribute to g_{form} shifts the CS−MT model prediction of the CMC for OSE closer to the experimental CMC value.

4.5. Modeling Results for Decyl Dimethyl Phosphine Oxide (C_{10}PO). CS−MT modeling results for the simulated C_{10}PO micelle are reported in Table 3. Theoretical predictions for the optimal micelles obtained using the CS−MT model and the traditional MT model, as well as the experimental data[39] for the micellization behavior of C_{10}PO in aqueous solution at 24 °C with 0.1 mM of added Na_2CO_3, are reported in Table 4. All the free-energy contributions, the g_{form} values, and the CMC values listed in Tables 3 and 4 were computed using the approach described in Section 4.2. In applying the CS−MT model and the traditional MT model, the surfactant head area, a_h, was modeled as being equal to 50 Å^2.[8]

Although computer simulation of C_{10}PO was done in a micelle with an aggregation number of 50, the optimal micelles predicted by the CS−MT model and the traditional MT model are small cylindrical micelles with a number-average aggregation number of 45. The predicted value of g_{int} ($4.08k_BT$) is slightly higher than that of $\hat{g}_{int}$ ($3.86k_BT$) because the optimal micelle is predicted to have a slightly larger aggregation number than that of the simulated micelle. For this surfactant, the predictions of

the CS−MT model and the traditional MT model for g_{form} and the CMC are quite similar. Our estimate of $g_{tr,CS-MT}$ ($-16.75k_BT$) is only $0.21k_BT$ less negative than our estimate of g_{tr} ($-16.96k_BT$), but nevertheless, this leads to a CMC prediction that is closer to the experimental CMC value than the CMC predicted using traditional MT modeling. In the CS−MT model, it is interesting to note that, despite their high f values, groups 1 and 2 (as defined in Figure 6) contribute a total of $-0.65k_BT$ to the hydrophobic driving force for micelle formation. Despite this negative free-energy contribution from the $C_{10}PO$ head, the CMC predicted by the CS−MT model is higher than the CMC predicted by the MT model. This can be explained by the fact that the CS−MT modeling approach models the hydration of the surfactant tail and its contribution to $g_{tr,CS-MT}$ in a different manner than the traditional MT model.

4.6. Modeling Results for Decyl Methyl Sulfoxide ($C_{10}SO$). CS−MT modeling results for the simulated $C_{10}SO$ micelle are reported in Table 3. Theoretical predictions for the optimal micelles obtained using the CS−MT model, as well as the traditional MT model and the experimental data[39] for the micellization behavior of $C_{10}SO$ in aqueous solution at 24 °C with 0.1 mM of added Na_2CO_3, are reported in Table 4. The results reported in Tables 3 and 4 for $C_{10}SO$ were computed using the approach described in Section 4.2. In applying the CS−MT model and the traditional MT model, the surfactant head area, a_h, was modeled as being equal to 30 Å^2.[8]

Like $C_{10}PO$, the $C_{10}SO$ micelle was simulated with an aggregation number of 50. However, the CS−MT model prediction for the optimal aggregation number was found to correspond to large cylindrical micelles with a number-average aggregation number of 12 802. The predicted value of g_{int} for $C_{10}SO$ ($2.94k_BT$) is significantly lower than the predicted value of g_{int} for $C_{10}PO$ ($4.08k_BT$). For $C_{10}SO$, the CS−MT model estimate of $g_{tr,CS-MT}$ ($-15.56k_BT$) is $1.4k_BT$ larger than the traditional MT model estimate of g_{tr} ($-16.96k_BT$), making the CS−MT model estimate of the CMC significantly larger than that of the traditional MT model and closer to the experimental CMC value. Group 1 of $C_{10}SO$ (see Figure 7) contributes $-0.66k_BT$ to $g_{tr,CS-MT}$, but the overall hydrophobic driving force for micelle formation is predicted to be higher using the CS−MT model than using the traditional MT model. In the next section, the CMC predicted by the CS−MT model and the traditional MT model for binary mixtures of $C_{10}PO$ and $C_{10}SO$ will be presented, and the similarities and differences between the various free-energy contributions to g_{form} for these two surfactants will be discussed in greater detail.

4.7. Modeling Results for Binary Mixtures of $C_{10}PO$ and $C_{10}SO$. In Figure 12, we present both CS−MT model predictions (see the — results) and traditional MT model predictions (- - -) for CMCs of binary mixtures of $C_{10}PO$ and $C_{10}SO$ in aqueous solution at 24 °C with 0.1 mM of added Na_2CO_3. Experimental mixture CMC data (◆) is presented for comparison.[39] CMC values are reported in mM on the y-axis, and the mole fraction of $C_{10}PO$ is reported on the x-axis. The mixture CMC values increase monotonically as the mole fraction of $C_{10}PO$ is increased. Both the CS−MT model and the traditional MT model capture this increase and yield reasonable estimates of the slope associated with this increase. However, the CMCs predicted by the CS−MT model are clearly in better agreement with the experimental CMC values than the CMCs predicted by the traditional MT model.

It is instructive to compare each of the free-energy contributions to g_{form} for $C_{10}PO$ and $C_{10}SO$, as estimated using the CS−MT model and the traditional MT model (see Tables 3 and 4).

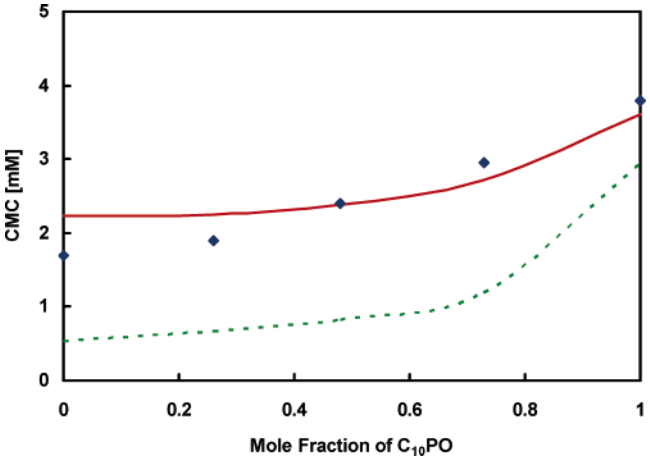


Figure 12. Predicted (— and - - -) and experimental (◆) CMCs of mixtures of $C_{10}PO$ and $C_{10}SO$, where the full line corresponds to CMCs predicted by the CS−MT model and the dashed line corresponds to CMCs predicted by the traditional MT model.

The discrepancy between predictions made using the CS−MT model and the traditional MT model can be explained in terms of the difference between the CS−MT model estimates of $g_{tr,CS-MT}$ and the traditional MT model estimates of g_{tr} for these two surfactants (see Table 3). For $C_{10}PO$, $g_{tr,CS-MT} = -16.75k_BT$ and $g_{tr} = -16.96k_BT$; for $C_{10}SO$, $g_{tr,CS-MT} = -15.56k_BT$ and $g_{tr} = -16.96k_BT$. As can be seen, the traditional MT model predictions for g_{tr} are the same for both $C_{10}PO$ and $C_{10}SO$, because both surfactants have the same number of hydrophobic groups in their tails. The CS−MT model predictions for the transfer free-energy contribution are less negative than those of the traditional MT model, making the CMC predicted using the CS−MT modeling approach higher for both surfactants and closer to the experimental CMC values (see Table 4).

Although the CS−MT model yields more accurate CMC predictions than the traditional MT model for $C_{10}PO$ and $C_{10}SO$, both the CS−MT model and the traditional MT model correctly predict that $C_{10}PO$ has a higher CMC than $C_{10}SO$. In the traditional MT model, the only molecular difference between $C_{10}PO$ and $C_{10}SO$ is their different head sizes (as captured in a_h). Therefore, the physical origin of the difference in CMC predictions for $C_{10}PO$ and $C_{10}SO$ made by the traditional MT model is due to steric effects (g_{st}). However, because the various free-energy contributions are coupled through the minimization of g_{form}, the influence of these steric effects is also manifested in the predicted values of g_{int} and g_{pack}. In the CS−MT modeling approach, the predicted CMCs also differ because of the different $g_{tr,CS-MT}$ values of $C_{10}PO$ and $C_{10}SO$. The steric effects arising from differences in head size, however, more than compensate for the more negative $g_{tr,CS-MT}$ value of $C_{10}PO$ ($-16.75k_BT$) relative to that of $C_{10}SO$ ($-15.56k_BT$) predicted using the CS−MT model, making the predicted CMC of $C_{10}PO$ 1.38 mM higher than that of $C_{10}SO$. In traditional MT modeling, g_{tr} for both surfactants is predicted to be the same, and therefore, the predicted CMC of $C_{10}PO$ is 2.39 mM larger than the predicted CMC of $C_{10}SO$. The experimental CMC difference for these two surfactants is 2.1 mM, which is closer to the CS−MT model prediction for the difference in the CMCs than to the traditional MT prediction for the difference in the CMCs.

4.8. Modeling Results for Dodecyl Octa(Ethylene Oxide) ($C_{12}E_8$). CS−MT modeling results for the simulated $C_{12}E_8$ micelle are reported in Table 3. Theoretical predictions for the optimal micelles obtained using the CS−MT model and the

traditional MT model, as well as the experimental data[40] for the micellization behavior of $C_{12}E_8$ in aqueous solution at 25 °C, are reported in Table 4. The same computational approach described in Section 4.2 was used to generate each of the free-energy contributions to g_{form} and the CMC values listed in Table 3, although for this surfactant, the CS−MT model predictions were made using two different approaches. In the first approach, reported under the heading "all hydr groups" in Table 3, the CS−MT model for g_{dehydr} (given in eq 2) was used to compute the free-energy contribution of every hydrophobic group in the surfactant molecule, regardless of whether the group is part of the surfactant head or tail. This is the modeling approach that has been used to model each of the other nonionic surfactants considered in this article. However, we also implemented CS−MT modeling of $C_{12}E_8$ in an alternate way, reported under the heading "tail hydr groups" in Table 3. In this alternate approach, only the hydrophobic groups in the surfactant tail (i.e., groups 26−37 listed in Figure 8) are included in the sum given in eq 2 to calculate g_{dehydr}. In both implementations of the CS−MT model, as well as in the implementation of the traditional MT model, the surfactant head area, a_h, was modeled as being equal to 53 Å^2.[8]

$C_{12}E_8$ differs from the other six nonionic surfactants considered in this article in that the traditional MT model CMC prediction (0.24 mM) and the experimental CMC (0.1 mM) are in reasonable agreement, but the "all hydr groups" CS−MT modeling approach severely underestimates the CMC (0.002 mM). The regular CS−MT modeling approach fails in this case because of the simplistic manner in which we have used eq 2 for $C_{12}E_8$. As discussed in Section 4.2, the values of g_{tr_i} used in eq 2 are strictly accurate only for the transfer of a hydrophobic oil group (CH, CH_2, or CH_3) from bulk water to a bulk phase of tails. For eq 2 to yield reasonable results, an accurate estimate of g_{tr_i} must first be made for each of the CH_2 groups in the E_8 head of $C_{12}E_8$. As discussed in Section 4.1, we have made the approximation that, for each of the nonionic surfactants modeled here, the g_{tr_i} values corresponding to hydrophobic groups in the surfactant head are equal to the g_{tr_i} values for the corresponding groups in the surfactant tail. For $C_{12}E_8$, however, we believe that this approximation is not sufficiently accurate for the following reasons: (i) using a water-to-oil transfer free energy for the process of transferring a CH_2 group in E_8 from bulk water to the corona region of a $C_{12}E_8$ micelle (which has a high concentration of water and other ethylene oxide groups) is a poor approximation, (ii) each of the hydrophobic groups in the surfactant head is bonded to a hydrophilic oxygen atom, thus affecting its hydrophobicity and g_{tr_i} value, and (iii) the large number of hydrophobic groups in the E_8 head (a total of 16) amplifies the effect of errors inherent in (i) and (ii) to a greater extent than that observed in the case of the other six nonionic surfactants considered in this article, which have relatively small, non-polymeric heads.

The CS−MT model tail hydr groups approach actually yields the most accurate prediction of the CMC when compared with the experimental CMC. Clearly, for $C_{12}E_8$, approximating each of the hydrophobic groups in the E_8 head of $C_{12}E_8$ as not contributing at all to the hydrophobic driving force for micelle formation is more appropriate than modeling them as contributing to the hydrophobic driving force with the same g_{tr_i} values as those corresponding to the CH_2 groups in the $C_{12}E_8$ tail. On the basis of the modeling results presented here, we conclude that care must be taken in applying the CS−MT model to surfactants with relatively long, polymeric heads. Without an accurate estimate of the appropriate g_{tr_i} values to use for the hydrophobic groups in the surfactant head, application of eq 2 to quantify the hydrophobic effect may not yield accurate results.

Although we do not explore these here, a number of approaches could be used to obtain more accurate estimates of g_{tr_i} values for the $C_{12}E_8$ head. Perhaps, the most straightforward approach would involve using an experimental or computational method to estimate the transfer free energy of an ethylene oxide monomer from bulk aqueous solution to a bulk phase of water and poly(ethylene oxide) molecules that serves as a reasonable proxy for the anisotropic corona region of the micelle. After obtaining the g_{tr_i} value of an ethylene oxide monomer, eq 2 could be used to calculate g_{dehydr}, albeit with the summation given in eq 2 extended to include all the hydrophobic groups in the $C_{12}E_8$ tail as well as all the ethylene oxide groups in the $C_{12}E_8$ head. To estimate g_{tr_i} of an ethylene oxide monomer, the solvation free energy of an ethylene oxide group in water and the solvation free energy of an ethylene oxide group in a bulk phase of water and poly(ethylene oxide) molecules could be determined experimentally. Alternatively, a theoretical approach could be used to estimate the transfer free energy or solvation free energy, for example, by using the Flory−Huggins approach with appropriate χ parameters or using a computer simulation approach that uses a realistic force field to describe interactions between the system components.[2,41]

4.9. Modeling Results for Decanoyl-*n*-Methylglucamide (MEGA-10). CS−MT modeling results for the simulated MEGA-10 micelle are reported in Table 3, and theoretical predictions made using the CS−MT model and the traditional MT model for the micellization behavior of MEGA-10 are reported in Table 4. Experimental data is also reported in Table 4 for comparison.[42] All data was generated for MEGA-10 in aqueous solution at 30 °C with 0.1 M of added NaCl. Each value reported in Tables 3 and 4 for MEGA-10 was computed using the approach described in Section 4.2. In applying the CS−MT model and the traditional MT model, the surfactant head area, a_h, was modeled as being equal to 62 Å^2.[8]

Computer simulation of MEGA-10 was conducted for a micelle with an aggregation number of 42. However, the optimal MEGA-10 micelle shape and size predicted by the CS−MT model are small spheres with a number-average aggregation number of only 22. As a result, the predicted value of g_{int} ($4.96k_BT$) is significantly higher than the predicted value of $\hat{g}_{int}$ ($3.81k_BT$). The low predicted aggregation number of MEGA-10 is due in part to its large head area, which at 62 Å^2 is larger than that of any of the other nonionic surfactants modeled here. This large head area, in turn, results in MEGA-10 having the highest value of g_{st} ($1.97k_BT$) among all the nonionic surfactants modeled. For MEGA-10, $g_{tr,CS-MT}$ ($-17.36k_BT$) is significantly more negative than g_{tr} ($-15.47k_BT$) computed using the traditional MT modeling approach. This is due primarily to the hydrophobic free-energy contributions calculated in the CS−MT modeling approach for groups 11 and 12 (see Figure 9). In the traditional MT modeling approach, the most reasonable estimate of the head and tail of MEGA-10 would be to include only groups 16−24 in the linear alkyl chain as part of the tail. On the basis of this tail assignment, the traditional MT model prediction of g_{tr} is less negative than the CS−MT model prediction of $g_{tr,CS-MT}$. As a result, the CMC predicted by the CS−MT model (6.55 mM) is much closer to the experimental CMC (5 mM) than the CMC predicted by the traditional MT model (43.33 mM). Clearly, only the CS−MT model predictions are reasonable for MEGA-10.

4.10. Effect of the Definition of Hydrating Contacts on the Modeling Results. In article 1, we stated that, in the context

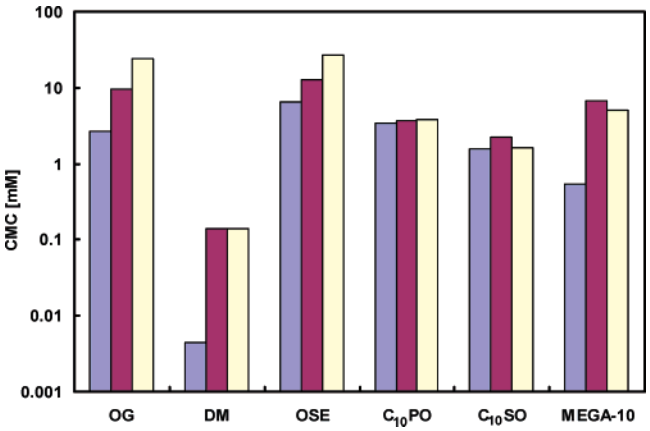


Figure 13. Comparison of CMCs predicted by the CS−MT model using f_{water} (blue) and f (maroon) values (see Section 4.10) in eqs 2 and 4 to compute g_{dehydr} and g_{hydr}, respectively. The definitions of f and f_{water} are given in Section 3.1. The experimental CMCs (yellow) are also reported for comparison with the theoretical predictions. The CMC values are reported on a log scale.

of CS−MT modeling, an atom in contact with a hydrophobic group is considered "hydrating" if the atom is capable of (i) forming hydrogen bonds or (ii) coordinate (dative covalent) bonding. On the basis of this definition, both water and hydrogen-bonding groups in surfactant heads contribute hydrating contacts when computing f, and therefore, one should use f values, rather than f_{water} values, in implementing the CS−MT modeling approach.

The effect of the way in which hydrating contacts are defined on our modeling results can be evaluated by comparing the CS−MT model predictions for the CMC using f values in eqs 2 and 4 with the CS−MT model predictions of the CMC using f_{water} values in eqs 2 and 4. Figure 13 compares the CMC predictions for OG, DM, OSE, $C_{10}PO$, $C_{10}SO$, and MEGA-10 using both approaches with the experimental CMC values. Results for $C_{12}E_8$ are not shown because the CS−MT model does not yield sufficiently accurate CMC predictions for this surfactant (for reasons discussed in Section 4.8). CMC results are reported on a log scale because of the large range spanned by the predicted and the experimental CMC values. With the exception of $C_{10}SO$, using f values (see the maroon results) in eqs 2 and 4 yields better CMC predictions than using f_{water} values (blue) in eqs 2 and 4. If only water molecules are considered to be hydrating, the degree of dehydration of the micelle core is overestimated, resulting in a prediction of g_{form} that is often significantly more negative than that predicted using the traditional MT model. As expected intuitively, the discrepancy between the f and f_{water} modeling results is largest for surfactants with many hydrogen-bonding groups in the surfactant head. For example, for OG, by defining hydration using f_{water} instead of using f, the predicted CMC is reduced by a factor of 3.6. For DM, defining hydration using f_{water} instead of using f has an even greater effect and reduces the predicted CMC by a factor of 31.

Modeling contacts with hydrogen-bonding groups in the surfactant head as being hydrating is consistent with the approximation made in traditional MT modeling that the surfactant heads at the micelle core−water interface shield only 21 Å^2 of the interface from hydrating contacts, an area corresponding to the cross-sectional area of the linear alkyl chain connected to the surfactant head, rather than to the cross-sectional area of the surfactant head. In traditional MT modeling, if the surfactant head is modeled as shielding its cross-sectional area, implying that the head itself does not provide hydrating contacts, less accurate predictions of micellization behavior would be obtained.

5. Conclusions

In this article, we have demonstrated the validity and accuracy of the CS−MT model by using it to model seven nonionic surfactants. To implement the CS−MT model, we have conducted two independent MD simulations for each nonionic surfactant to determine information about the changes in hydration that occur upon micelle formation. Changes in hydration were quantified by computing a fractional hydration value, f, for each group. The f values obtained for each surfactant through MD simulation were used as an input in a new model presented in article 1 of this series that computes the magnitude of the hydrophobic free-energy contribution as the sum of g_{dehydr} and g_{hydr}. In this article, we have used a simple computational strategy to estimate g_{dehydr} and g_{hydr} for each nonionic surfactant modeled. To calculate g_{dehydr}, g_{tr_i} values were estimated using the same solubility correlations for linear alkyl tails that our group has used in the past in traditional MT modeling.[32] To calculate g_{hydr}, we have made the simple approximation that this contribution is equal to zero for any hydrophobic group with a value of f greater than 0.60. For hydrophobic groups with a value of f less than 0.60, g_{hydr} was calculated using eq 4 and an expression for Δg_{wc} derived for oil molecules in article 1 of this series. By combining elements of the CS−MT model and the traditional MT model, the hydrophobic driving force for micelle formation was quantified as a transfer free energy ($g_{tr,CS-MT}$). After determining this input, the free energy of micelle formation, g_{form}, and the CMC were calculated for each surfactant for micelles of the optimal shape and size.

Reasonable agreement between the CS−MT model predictions and the experimental data for g_{form} and the CMC were obtained for OG, DM, OSE, $C_{10}PO$, $C_{10}SO$, and MEGA-10. For $C_{12}E_8$, the CS−MT model predictions were not in good agreement with the experimental data because the simple approximations that were made in this article to estimate g_{dehydr} were not sufficiently accurate for this surfactant. For the 16 hydrophobic CH_2 groups in the E_8 head of $C_{12}E_8$, we believe that more accurate values of g_{tr_i} must be used to obtain accurate modeling results. Consequently, we recommend that the simple approximations that we have made to compute g_{dehydr} only be used for surfactants with relatively small, non-polymeric heads.

The predictions of the CS−MT and the traditional MT modeling approaches were found to be in reasonable agreement for OG, DM, OSE, $C_{10}PO$, and $C_{10}SO$. For four of these surfactants (DM, OSE, $C_{10}PO$, and $C_{10}SO$), the CMCs predicted by the CS−MT model were closer to the experimental CMC values than the CMCs predicted by the traditional MT model. In addition, CMCs predicted for binary mixtures of $C_{10}PO$ and $C_{10}SO$ using the CS−MT modeling approach were significantly closer to the experimental CMCs than those predicted using the traditional MT modeling approach. The CMC predicted by the CS−MT model for MEGA-10 (6.55 mM) was significantly closer to the experimental CMC value (5 mM) than the CMC predicted by the traditional MT model (43.3 mM).

Using the f values obtained through computer simulation, we evaluated the accuracy of two approximations made in traditional MT modeling to quantify the hydrophobic driving force for micelle formation in aqueous solution for the first time. These include (i) the accuracy of using surfactant head and tail assignments to compute g_{tr} and (ii) the extent to which surfactant heads shield the micelle core from hydrating contacts. Approximation (i) was found to be reasonable for nonionic

surfactants in the absence of detailed hydration data. Approximation (ii) was found to be accurate for some of the nonionic surfactants modeled in this article but not very accurate for others. However, the approximation made in the traditional MT model that a surfactant head shields 21 Å^2 at the micelle core−water interface from hydrating contacts was found to be in close agreement with the average computer simulation estimate for this shielded area (21.85 Å^2). An important advantage of the CS−MT modeling approach is that it eliminates the need to make approximations (i) and (ii) by enabling determination of the changes in hydration that occur during micelle formation, as well as providing a theoretical model to quantify the hydrophobic driving force for micelle formation in aqueous solution directly from this hydration data.

Using the relatively simple approach to estimate g_{dehydr} and g_{hydr} presented in this article, the CS−MT model was found to yield similar, or superior, predictions of the CMCs of nonionic surfactants with relatively small, non-polymeric heads when compared to the traditional MT model. The results obtained for the relatively complex surfactant MEGA-10 highlight the strengths of the CS−MT modeling approach: for surfactants where it is difficult to make accurate head and tail assignments for traditional MT modeling and for surfactants where a significant number of hydrophobic groups are located near the aggregate core−water interface and are partially hydrated, the CS−MT modeling approach is expected to yield better results than the traditional MT modeling approach. Clearly, the more accurate predictions of the CS−MT model come at the cost of greater computational expense. Nevertheless, given the relatively small fraction of surfactants with sufficient structural and chemical simplicity to be easily modeled using the traditional MT model, we conclude that the CS−MT modeling approach represents a very promising alternative.

Acknowledgment. This research was supported in part by funding provided by DuPont through the DuPont−MIT Alliance. B.C.S. would like to thank Kate Stafford for her assistance in setting up several of the nonionic surfactant simulations. B.C.S. and A.G. are also grateful to Dr. Isaac Reif for critically reading the manuscript and for stimulating discussions about the CS−MT model.

Supporting Information Available: Two figures are available that plot SASA normalized by the average value of SASA as a function of simulation time during the 5 ns data-gathering simulation run (see Section 2.3) for each of the seven nonionic surfactant micelles simulated in this article. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Stephenson, B. C.; Goldsipe, A.; Beers, K. J.; Blankschtein, D. *J. Phys. Chem. B* **2007**, *111*, 1025−1044.
(2) Nagarajan, R.; Ruckenstein, E. *Langmuir* **1991**, *7*, 2934−2969.
(3) Shiloach, A.; Blankschtein, D. *Langmuir* **1998**, *14*, 1618−1636, and references cited therein.
(4) Gunnarsson, G.; Jonsson, B.; Wennerstrom, H. *J. Phys. Chem.* **1980**, *84*, 3114−3121.
(5) Jonsson, B.; Wennerstrom, H. *J. Colloid Interface Sci.* **1981**, *80*, 482−496.
(6) Evans, D. F.; Mitchell, D. J.; Ninham, B. W. *J. Phys. Chem.* **1984**, *88*, 6344−6348.
(7) Hayter, J. B. *Langmuir* **1992**, *8*, 2873−2876.
(8) Puvvada, S.; Blankschtein, D. *J. Chem. Phys.* **1990**, *92*, 3710−3724, and references cited therein.
(9) van der Spoel, D.; Lindahl, E.; Hess, B.; van Buuren, A. R,; Apol, E.; Meulenhoff, P. J.; Tieleman, D. P.; Sijbers, A. L. T. M.; Feenstra, K. A.; van Drunen, R.; Berendsen, H. J. C. *Gromacs User Manual version 3.2*; www.gromacs.org:2004.
(10) Goldsipe, A.; Blankschtein, D. *Langmuir* **2005**, *22*, 9850−9865, and references cited therein.
(11) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.
(12) Chirlian, L. E.; Francl, M. M. *J. Comput. Chem.* **1987**, *8*, 894−905.
(13) Stephenson, B. C.; Beers, K.; Blankschtein, D. *Langmuir* **2006**, *22*, 1500−1513, and references cited therein.
(14) Stephenson, B. C.; Rangel-Yagui, C. O.; Pessoa, A.; Tavares, L. C.; Beers, K.; Blankschtein, D. *Langmuir* **2006**, *22*, 1514−1525.
(15) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(16) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(17) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1997**, *23*, 327−341.
(18) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. *Comput. Phys. Commun.* **1995**, *91*, 43−56.
(19) Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Model.* **2001**, *7*, 306−317.
(20) Jang, S. S.; Goddard, W. A., III *J. Phys. Chem. B* **2006**, *110*, 7992−8001.
(21) Stephenson, B. C.; Beers, K. J. *J. Phys. Chem. B* **2006**, *110*, 19393−19405.
(22) Feller, S.; Pastor, R. *Biophys. J.* **1996**, *71*, 1350−1355.
(23) Israelachvili, J. N. *Intermolecular and Surface Forces*, 2nd ed.; Academic Press: New York, 1991.
(24) Bruce, C.; Berkowitz, M.; Perera, L.; Forbes, M. D. E. *J. Phys. Chem. B* **2002**, *106*, 3788−3793.
(25) Flyvbjerg, H.; Petersen, H. G. *J. Chem. Phys.* **1989**, *91*, 461−466.
(26) Hess, B. Thesis. *Stochastic Concepts in Molecular Simulation*; Rijksuniversiteit Groningen, 1999.
(27) Hess, B. *J. Chem. Phys.* **2002**, *116*, 209−217.
(28) Gibbs, J. W. *The Scientific Papers of J. W. Gibbs*; Dover: New York, 1961; Vol. 1.
(29) Koenig, F. O. *J. Chem. Phys.* **1950**, *18*, 449.
(30) Buff, F. P. *J. Chem. Phys.* **1951**, *19*, 1591.
(31) Tolman, R. C. *J. Chem. Phys.* **1948**, *16*, 758.
(32) Smith, R.; Tanford, C. *Proc. Natl. Acad. Sci.* **1973**, *70*, 289−293.
(33) Reynolds, J. A.; Gilbert, D. B.; Tanford, C. *Proc. Natl. Acad. Sci.* **1974**, *71*, 2925−2927.
(34) Stephenson, B. C.; Beers, K. J.; Blankschtein, D. *J. Phys. Chem. B* **2007**, *111*, 1063−1075.
(35) Mukerjee, P.; Chan, C. C. *Langmuir* **2002**, *18*, 5375−5381.
(36) Srinivasan, V.; Blankschtein, D. *Langmuir* **2003**, *19*, 9932−9945, and references cited therein.
(37) Rosen, M. J.; Sulthana, S. B. *J. Colloid Interface Sci.* **2001**, *239*, 528−534.
(38) Aratono, M.; Kanda, T.; Motomura, K. *Langmuir* **1990**, *6*, 843−846.
(39) Holland, P. M.; Rubingh, D. N. *J. Phys. Chem.* **1983**, *87*, 1984−1990.
(40) Alargova, R. G.; Kochijashky, I. I.; Sierra, M. L.; Zana, R. *Langmuir* **1998**, *14*, 5412−5418.
(41) Carale, T. R.; Pham, Q. T.; Blankschtein, D. *Langmuir* **1994**, *10*, 109−121.
(42) Hierrezuelo, J. M.; Aguiar, J.; Ruiz, C. C. *Langmuir* **2004**, *20*, 10419−10426.

Quantifying the Hydrophobic Effect. 3. A Computer Simulation−Molecular-Thermodynamic Model for the Micellization of Ionic and Zwitterionic Surfactants in Aqueous Solution

Brian C. Stephenson, Kenneth J. Beers, and Daniel Blankschtein*

Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received: September 1, 2006; In Final Form: November 21, 2006

In this article, the validity and accuracy of the CS−MT model introduced in article 1 for oil aggregates and in article 2 for nonionic surfactants is further evaluated by using it to model the micellization behavior of ionic and zwitterionic surfactants in aqueous solution. In the CS−MT model, two separate free-energy contributions to the hydrophobic driving force for micelle formation are computed using hydration data obtained from computer simulation: g_{dehydr}, the free-energy change associated with dehydration, and g_{hydr}, the change in the hydration free energy. To enable straightforward estimation of g_{dehydr} and g_{hydr} for ionic and zwitterionic surfactants, a number of simplifying approximations were made. Reasonable agreement between the CMCs predicted using the CS−MT model and the experimental CMCs was obtained for sodium dodecyl sulfate (SDS), dodecylphophocholine (DPC), cetyltrimethylammonium bromide (CTAB), two 3-hydroxy sulfonate surfactants (AOS-12 and AOS-16), and a homologous series of four DC_NA bromide surfactants with a dimethylammonium head attached to a dodecyl alkyl tail and to an alkyl side chain of length C_N, having the chemical formula $C_{12}H_{25}C_NH_{2N+1}N(CH_3)_2Br$, with $N = 1$ (DC_1AB), 2 (DC_2AB), 4 (DC_4AB), and 6 (DC_6AB). For six of these nine surfactants, the CMCs predicted using the CS−MT model are closer to the experimental CMCs than the CMCs predicted using the traditional molecular-thermodynamic (MT) model. For DC_2AB, DC_4AB, and DC_6AB, which are the most structurally complex of the ionic surfactants modeled, the CMCs predicted using the CS−MT model are in remarkably good agreement with the experimental CMCs, and the CMCs predicted using the traditional MT model are quite inaccurate. Our results suggest that the CS−MT model accurately quantifies the hydrophobic driving force for micelle formation for ionic and zwitterionic surfactants in aqueous solution. For complex ionic and zwitterionic surfactants where it is difficult to accurately quantify the hydrophobic driving force for micelle formation using the traditional MT modeling approach, the CS−MT model represents a very promising alternative.

1. Introduction

The CS−MT model combines hydration data obtained through computer simulation with a free-energy model for the hydrophobic driving force for micelle formation to model surfactant micellization and micellar solubilization in aqueous solution. This new modeling approach was introduced in article 1 of this series and tested by modeling oil aggregates.[1] The primary motivation for the development of the CS−MT model is to extend the traditional MT model to increasingly chemically and structurally complex surfactants and solubilizates. In article 2 of this series,[2] the CS−MT model was extended to model the micellization behavior of nonionic surfactants by outlining a simple computational strategy to estimate g_{dehydr}, the free-energy change associated with dehydration, and g_{hydr}, the change in the hydration free energy, when non-charged hydrophilic head groups are present at the micelle core−water interface. For each surfactant, g_{dehydr} was calculated by estimating g_{tr_i} values using solubility correlations for linear alkyl tails.[5] The approximation was made that g_{tr_i} values could be estimated in this manner for hydrophobic groups in both the surfactant head and tail.[2] The free-energy contribution, g_{hydr}, was calculated for each surfactant by (i) approximating g_{hydr} as being equal to zero for any hydrophobic group that is not adsorbed onto, or incorporated within, the micelle core and (ii) estimating g_{hydr} for hydrophobic groups in the micelle core using the expression for Δg_{wc} given in eqs 5 and 6. As justified in article 2, any hydrophobic group with a value of f greater than 0.60 was not considered to be part of the micelle core. In computing Δg_{wc} using eqs 5 and 6, the approximation is made that the change in hydration free energy experienced by hydrophobic groups as they are transferred from the bulk aqueous solution to the micelle core is unaffected by the presence of the surfactant heads at the micelle core−water interface. In article 2, the CS−MT model was used to predict the micellization behavior of seven nonionic surfactants with varying degrees of structural complexity. In this article, we use the CS−MT model to predict the micellization behavior of ionic and zwitterionic surfactants that are both simple and challenging to model using the traditional MT modeling approach. The most complex of the surfactants considered in this article are too complex to model accurately using the traditional MT model.

1.1. Overview of the CS−MT Model. The free energy of aggregate formation, g_{form}, is computed in the CS−MT model as the sum of the following six free-energy contributions:[1,2]

$$g_{form} = g_{dehydr} + g_{hydr} + g_{pack} + g_{st} + g_{elec} + g_{ent} \quad (1)$$

* Corresponding author. Department of Chemical Engineering, Room 66-444 Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139. Telephone: (617) 253-4594. Fax: (617) 252-1651. E-mail: dblank@mit.edu.

10.1021/jp065699v CCC: $37.00
Published on Web 01/18/2007

Two of the free-energy contributions to g_{form}, g_{dehydr} and g_{hydr}, reflect the hydrophobic free-energy change associated with aggregate formation or the hydrophobic driving force for aggregate formation. In the CS−MT modeling approach, computer simulation data on surfactant hydration in the bulk water and in the micellar states is used to compute both g_{dehydr} and g_{hydr}. They are computed using the following two equations:[1,2]

$$g_{dehydr} = \sum_{i=1}^{n_{hydr}} (1 - f_i) g_{tr_i} \quad (2)$$

$$g_{hydr} = \sum_{i=1}^{n_{core}} SASA_i f_i \Delta g_{wc_i} \quad (3)$$

where n_{hydr} is the total number of hydrophobic groups in the solute, f_i is the fractional hydration of group i, g_{tr_i} is the free-energy change associated with transferring group i from the aqueous solution to a bulk solution composed of solute tails, n_{core} is the total number of hydrophobic groups in the solute that adsorb onto, or penetrate into, the aggregate core, $SASA_i$ is the solvent accessible surface area of group i, and Δg_{wc_i} is defined as the difference in the free energy per unit of solvent accessible surface area associated with the hydration of group i in the micellar state and the aqueous solution. In articles 1 and 2, we justified computing f for each group i as follows:[1,2]

$$f = \frac{\text{number of hydrating contacts in the aggregate}}{\text{number of hydrating contacts in bulk water}} \quad (4)$$

where a "hydrating contact" is defined as a contact with an atom that (i) hydrogen bonds or (ii) is capable of coordinate (dative covalent) bonding. In article 1, we justified the use of a 0.3 nm cutoff distance to count the hydrating contacts that occur during MD simulation. This cutoff distance was shown to be appropriate in the context of modeling nonionic surfactant micellization in article 2. In article 1, we developed a theoretical model to estimate Δg_{wc_i} for oil molecules. For oil molecules, Δg_{wc_i} does not depend on i and is given by[1]

$$\Delta g_{wc} = \sigma_{core} - \sigma_{bulk} = \frac{\sigma A_{core}}{SASA_{core}} - \frac{g_{tr_i}}{SASA_i} \quad (5)$$

where σ_{core} is the microscopic "interfacial tension" (interfacial free energy per unit SASA) associated with the aggregate core−water interface, σ_{bulk} is the microscopic "interfacial tension" (interfacial free energy per unit SASA) associated with the group i (CH_2 or CH_3)−water interface in the aqueous solution, σ is the macroscopic interfacial tension of the aggregate core−water interface, A_{core} is the area of the hydrophobic aggregate core as computed geometrically from the volume of the aggregate and the assumption of a perfectly smooth aggregate surface, and $SASA_{core}$ is the solvent accessible surface area of the hydrophobic aggregate core. The ratio $A_{core}/SASA_{core}$ in eq 5 was estimated using the following correlation:[1]

$$SASA_{core}/A_{core} = 1.740 - 0.026 n_t + 0.078 C \quad (6)$$

where n_t is the total number of hydrophobic groups in the solute that are part of the hydrophobic aggregate core and C is the curvature of the micellar aggregate, which is defined as $2/l_c$ for spheres, $1/l_c$ for cylinders, and zero for planar interfaces, where l_c is the core-minor radius or planar half-width. In article 2, we demonstrated that the model for Δg_{wc} given by eqs 5 and 6 can be used to model the change in hydration free energy experienced by the hydrophobic CH, CH_2, and CH_3 groups present in the hydrophobic core of a nonionic surfactant micelle.

We note that the remaining four free-energy contributions appearing in eq 1 (g_{pack}, g_{st}, g_{elec}, and g_{ent}) are computed in the CS−MT model in the same way that they are computed in the traditional MT modeling approach.[3] However, as noted in article 1, the way in which g_{pack} and g_{st} are computed may be informed by molecular dynamics simulation data. With this in mind, in this article, we explore an approach to use computer simulation data to accurately estimate g_{st} for surfactants with complex head structures.

In Appendix A of article 1, we showed that by combining elements of the CS−MT model and the traditional MT model, g_{form} can be computed as a function of aggregate shape and size after only two computer simulations—one simulation of the solute in bulk water and one simulation of the solute in an aggregate of arbitrary shape and size. The free energy of aggregate formation, g_{form}, for micelles of a different shape and size than the simulated micelle is computed using the following equation:[1]

$$g_{form} = g_{tr,CS-MT} + g_{int} + g_{pack} + g_{st} + g_{elec} + g_{ent} \quad (7)$$

where $g_{tr,CS-MT}$ is the transfer free-energy contribution obtained using the CS−MT modeling approach. The term $g_{tr,CS-MT}$ is computed using the relationship $g_{tr,CS-MT} = g_{dehydr} + g_{hydr} - \hat{g}_{int}$, where $\hat{g}_{int}$ is the traditional MT model prediction for the interfacial free-energy contribution of the *simulated* micellar aggregate.

As discussed in articles 1 and 2, for a micelle of the optimum shape, size, composition, and degree of counterion binding, g_{form} has a minimum value, which we denote as g^*_{form}. By determining g^*_{form}, the optimal aggregate shape (S^*), the optimal core-minor radius (l^*_c), the optimal composition (α^*), and the optimal degree of counterion binding (β^*) can be predicted. In addition, the CMC in mole fraction units is computed as follows:[4]

$$CMC \approx \exp\left(\frac{g^*_{form}(S^*, l^*_c, \alpha^*, \beta^*)}{k_B T}\right) \quad (8)$$

where k_B is the Boltzmann constant and T is the absolute temperature.

1.2. Modeling Ionic and Zwitterionic Surfactant Micellization. In this article, we use the CS−MT modeling approach to model the micellization behavior of nine ionic and zwitterionic surfactants in aqueous solution. The surfactants selected for modeling include three anionic surfactants (sodium dodecyl sulfate (SDS) and two 3-hydroxy sulfonate surfactants with different hydrophobic tail lengths (AOS-12 and AOS-16)), one zwitterionic surfactant (dodecylphosphocholine (DPC)), and five cationic surfactants (cetyltrimethylammonium bromide (CTAB) and a homologous series of four DC_NA bromide surfactants with a positively charged dimethylammonium head attached to a dodecyl tail and to an alkyl side chain of length C_N, having the chemical formula $C_{12}H_{25}C_NH_{2N+1}N(CH_3)_2Br$, where $N = 1$ for what we will refer to as DC_1AB, $N = 2$ for DC_2AB, $N = 4$ for DC_4AB, and $N = 6$ for DC_6AB).

To use the CS−MT modeling approach to model ionic and zwitterionic surfactant micellization, a number of approximations must be made to account for the presence of charged or dipolar hydrophilic groups at the micelle core−water interface. The validity of the approximations proposed in article 2 to model

nonionic surfactant micellization with the CS−MT model is discussed and evaluated in this article for ionic and zwitterionic surfactant micellization. The CS−MT model is first used to predict the micellization behavior of a simple anionic surfactant (SDS), a simple zwitterionic surfactant (DPC), and a simple cationic surfactant (CTAB). The CMCs predicted using the CS-MT model for these surfactants are compared with the CMCs predicted using the traditional MT model as well as with the experimental CMCs taken from the literature. Based on the modeling results, we discuss whether the theoretical approach used to derive eqs 5 and 6 can be used to calculate Δg_{wc} in the case of ionic and zwitterionic surfactants or whether Δg_{wc} must be fit for these surfactants.

After determining the applicability of the CS−MT model for three simple ionic and zwitterionic surfactants, we use it to model six complex ionic surfactants (AOS-12, AOS-16, and the homologous series of DC_NAB surfactants). The CMCs predicted using the CS−MT model for these six surfactants are compared with the CMCs predicted using the traditional MT model as well as with the experimental CMCs taken from the literature.

The remainder of this article is organized as follows. Section 2 describes the computer simulation approach used in this article, including an overview of the modeling approach (Section 2.1), the simulation methods and parameters (Section 2.2), a description of how each system was prepared and equilibrated (Section 2.3), and the data analysis method used to analyze the molecular dynamics trajectories (Section 2.4). Computer simulation results are presented in Section 3. The CS−MT model is used to predict the micellization behavior of the nine surfactants considered in Section 4. Concluding remarks are presented in Section 5.

2. Molecular Dynamics Simulations

2.1. Modeling Approach. As discussed in articles 1 and 2, the CS−MT model requires fractional hydration data as an input. This fractional hydration data is obtained by performing two simulations. The first simulation is of a single surfactant molecule in a simulation cell of water (the "bulk water" simulation), and the second simulation is of the same surfactant molecule in a micellar environment (the "aggregate" simulation). As shown in Appendix A of article 1, obtaining information about the hydration state of a surfactant molecule in a micelle of a single shape and size is sufficient to allow prediction of the *optimal* micelle shape and size.

2.2. Simulation Methods and Parameters. The simulation methods and parameters used in this study are identical to those introduced in articles 1 and 2.[1,2,11−15] Surfactants were modeled using the fully atomistic OPLS-AA force field,[6] and water molecules were modeled using the simple extended point-charge (SPC/E) model for water. Some additional parameters to describe angles and angle vibrations were taken from the literature to model the sulfate (SO_4^-) group in SDS.[7] DPC was modeled using the GROMACS force field using the same parameters that were used by Tieleman et al.[8] The GROMACS force field models methylene (CH_2) groups and methyl (CH_3) groups as unified atoms.[9]

The atoms in the surfactant head for each ionic surfactant considered did not have suggested charges in the OPLS-AA force field, so the atomic charges for the heads of these surfactants were estimated using the CHelpG algorithm (as implemented in Gaussian 98).[10] We note that CHelpG was not used to assign atomic charges for the hydrophobic tail of each surfactant for reasons discussed in article 2. The atomic charges used by Tieleman et al. were used in modeling DPC.

TABLE 1: Overview of the MD Simulations[a]

surfactant	number of surfactant molecules	number of water molecules	total number of atoms
SDS	44	3347	11933
DPC	48	10453	10453
CTAB	49	10304	33999
AOS-12	32	3555	12041
AOS-16	40	6119	20557
DC_1AB	25	3522	11841
DC_2AB	20	3771	12393
DC_4AB	18	3734	12282
DC_6AB	16	3711	12189

[a] The geometry, the number of surfactant and water molecules, and the total number of atoms corresponding to each of the ionic and zwitterionic surfactant micelle simulations.

2.3. System Preparation and Equilibration. *2.3.1. Bulk Water Simulation.* The bulk water simulation for each surfactant considered was initialized by placing a single surfactant molecule in a simulation cell and surrounding it with sufficient water molecules to approximation infinite dilution (see articles 1 and 2). A single ion (which we will refer to hereafter as the counterion) was introduced into the simulation cell by replacing the water molecule experiencing the greatest electrostatic potential after initial energy minimization. After a brief *NPT* simulation to equilibrate the cell volume, data was gathered over the course of an additional 3 ns of *NPT* simulation.

2.3.2. Aggregate Simulation. Each surfactant micelle was preformed as a spherical aggregate in the manner discussed in article 2. To preserve electroneutrality, equal numbers of counterions and ionic surfactant molecules were added to each simulation cell. To speed equilibration, counterions were added by replacing water molecules experiencing the greatest electrostatic potential after initial energy minimization, with the potential being recalculated after every counterion insertion. In Table 1, we report the number of surfactant and water molecules and the total number of atoms included in each simulation cell.

Each micelle was built with a sufficiently small aggregation number to ensure that it would be spherical during simulation by estimating the expected spherical aggregation number based on the head area and tail volume of each surfactant.[16] As shown in Appendix A of article 1, a micelle of any aggregation number and geometry (whether spherical, cylindrical, or planar) may be simulated to obtain hydration information for the CS−MT modeling approach. However, as discussed in article 2, it is most convenient computationally to simulate spherical aggregates because this removes the need to either (i) specify a physically realistic area for each surfactant molecule within a cylindrical micelle or bilayer or (ii) specify a poorly defined interfacial tension perpendicular to the axis of a cylindrical micelle or the plane of a bilayer.

After preforming each spherical micelle, a 15 ns equilibration run was conducted under *NPT* conditions. For reasons discussed in article 2, we believe that this equilibration time is more than adequate to thoroughly equilibrate each micelle.[17] Plots of the equilibration profile of the SASA for three representative surfactants (SDS, DPC, and AOS-12) are shown in Figure 1. The SASA values reported in Figure 1 for each surfactant were normalized by the average value of SASA for that surfactant to facilitate comparison of the results. SASA was computed using a probe sphere of radius 0.2 nm (for a justification of this choice, see article 1) and using the double cubic lattice method as implemented in GROMACS.[13] Because obtaining accurate hydration information is the primary objective of our computer simulations, we consider SASA to be an important metric to evaluate equilibration, because this property is directly propor-

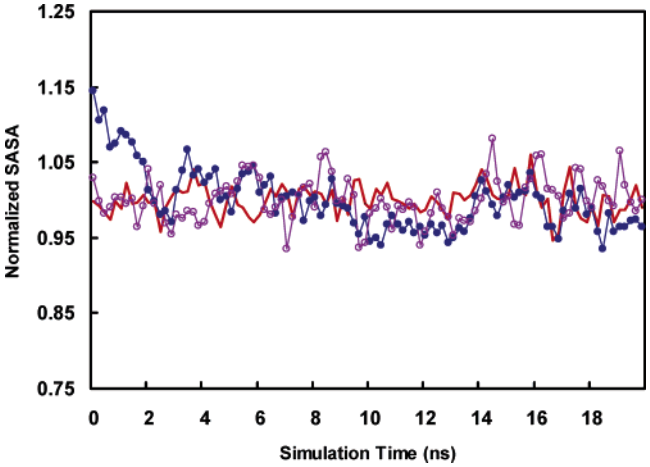


Figure 1. Solvent accessible surface area (SASA) normalized by the average value of SASA as a function of simulation time for micelles of three representative ionic and zwitterionic surfactants: sodium dodecyl sulfate (SDS, −), dodecylphosphocholine (DPC, −●−), and 3-hydroxy dodecyl sulfonate (AOS-12, −○−).

tional to the degree of hydration of the micelle. As shown in Figure 1, after the first 5 ns of simulation, there is no apparent drift in SASA for these three surfactants. Plots of the normalized values of SASA during the 5 ns data-gathering simulation runs for the remaining six surfactants considered are presented in the Supporting Information accompanying this article. All the equilibration SASA profiles indicate that the hydration data obtained during the data-gathering simulations are representative of the equilibrium hydration state of the micelle.

Snapshots of the post-equilibration configurations of each simulated ionic and zwitterionic surfactant micelle are shown in Figure 2 (water molecules and counterions have been omitted for clarity). Each surfactant molecule is depicted using the van der Waals radius of each atom.

2.4. Data Analysis Method. The fractional hydration of each atom (or group of atoms) in the surfactant molecule during the bulk water simulation and the aggregate simulation was quantified by counting hydrating contacts using a cutoff distance of 0.3 nm and by determining f from the contact data using eq 4.[1,2] For the ionic surfactants simulated in this article, contacts with water atoms, with hydrogen-bonding surfactant heads, and with counterions were each included as contributing to hydration. For the zwitterionic surfactant DPC, only contacts with water atoms and with hydrogen-bonding surfactant heads were counted as being hydrating, because no counterions were present in the simulation cell.

Although a cutoff of 0.3 nm was used in determining the f values for the CS−MT model, we note that to obtain estimates of f for some of the large atoms present in some of the surfactant heads considered (including nitrogen, sulfur, and oxygen), it was necessary to use a larger cutoff of 0.5 nm in order to obtain good statistics for f. Accordingly, a cutoff of 0.5 nm was used to generate all the hydration plots presented in Section 3.

An estimate of the standard error in f for each group of atoms in the surfactant molecule was obtained through the use of block averaging (for details, see article 1).[18−20] Data-gathering simulation runs were conducted for sufficient time to ensure that the uncertainty in each calculated value of f was small (typically less than 5%).

3. Simulation Results and Discussion

3.1. Sodium Dodecyl Sulfate (SDS). SDS is a widely used and extensively studied anionic surfactant. The fractional degree of hydration, f, of SDS is plotted as a function of group number in Figure 3. Groups in SDS that are considered to be part of the SDS head in traditional MT modeling (groups 1 and 2) have f values that are much larger than the f values of groups in the SDS tail (groups 3−13). However, even groups in the SDS head are partially dehydrated, with group 1 having an f value of 0.89 and group 2 having an f value of 0.63. The average f value of the groups in the SDS tail is 0.24.

The f results shown in Figure 3 reveal that the degree of dehydration of the groups in the SDS tail is a function of their distance from the SDS head. For example, group 3 in the tail (closest to the head) has an f value of 0.45, and group 13 (furthest from the head) has an f value of 0.18. We note that each of the groups in the SDS head has an f value greater than 0.60, and each of the groups in the SDS tail has an f value less than 0.60. As discussed in article 2, only hydrophobic surfactant groups with an f value below 0.60 are considered to be part of the micelle core and to have a nonzero value of g_{hydr}.

3.2. Dodecylphosphocholine (DPC). DPC is a zwitterionic surfactant frequently used as a model membrane lipid to study lipid-bound peptides and proteins.[21] DPC has been widely studied both experimentally and through computer simulations.[8,21] As discussed in Section 2.2, the force field parameters used to model this surfactant are the same as those used by Tieleman et al.[8,21,22]

The fractional degree of hydration, f, of DPC is plotted as a function of group number in Figure 4. The average f value of the groups in the DPC head (groups 1−8) is 0.70, and that of the groups in the DPC tail (groups 9−19) is 0.18. It is interesting to note that most of the hydrophobic groups in the DPC head (1−3, 5, and 6) are very hydrated, with f values of 0.69 or greater. As discussed in article 2, such groups do not contribute to g_{hydr} because they are not incorporated into the micelle core. However, they do contribute significantly to g_{dehydr}. It is also interesting to note that the hydrophobic CH_2 group adjacent to the DPC tail (group 8) has an f value of 0.40, which, as discussed in article 2, is sufficiently low that it is modeled as being part of the micelle hydrophobic core in CS−MT modeling. The low f value for this group is closer to the f value that would be expected if the group were adjacent to a nonionic surfactant head (for details, see article 2) than to an ionic surfactant head (see the f values of group 2 in SDS and of group 5 in CTAB).

3.3. Cetyltrimethylammonium Bromide (CTAB). CTAB is a commonly used cationic surfactant for which extensive experimental micellization data is available. The fractional degree of hydration, f, of CTAB is plotted as a function of group number in Figure 5. The average f value of the groups in the CTAB head (groups 1−5) is 0.85, and that of the groups in the CTAB tail (groups 6−20) is 0.20. As in the case of SDS and DPC, the hydrophobic groups in the CTAB head (1−3 and 5) are highly hydrated, with f values of 0.71 or greater. In general, the hydration profile observed for the anionic surfactant SDS is similar to the hydration profile observed for the cationic surfactant CTAB. For example, the average f value of the groups in the SDS head is 0.85, and the average f value of the groups in the CTAB head is also 0.85. In addition, the first CH_2 group in the SDS tail (group 3) and the CTAB tail (group 6) have f values of 0.45 and 0.57, respectively. In contrast, the zwitterionic surfactant DPC exhibits significant differences in the hydration of its head groups relative to those in SDS and in CTAB (see Section 3.2).

3.4. Sodium 3-Hydroxy Sulfonates (AOS-12 and AOS-16). Sodium α-olefinsulfonates (AOS) are anionic surfactants frequently used in household and industrial formulations.[23] These

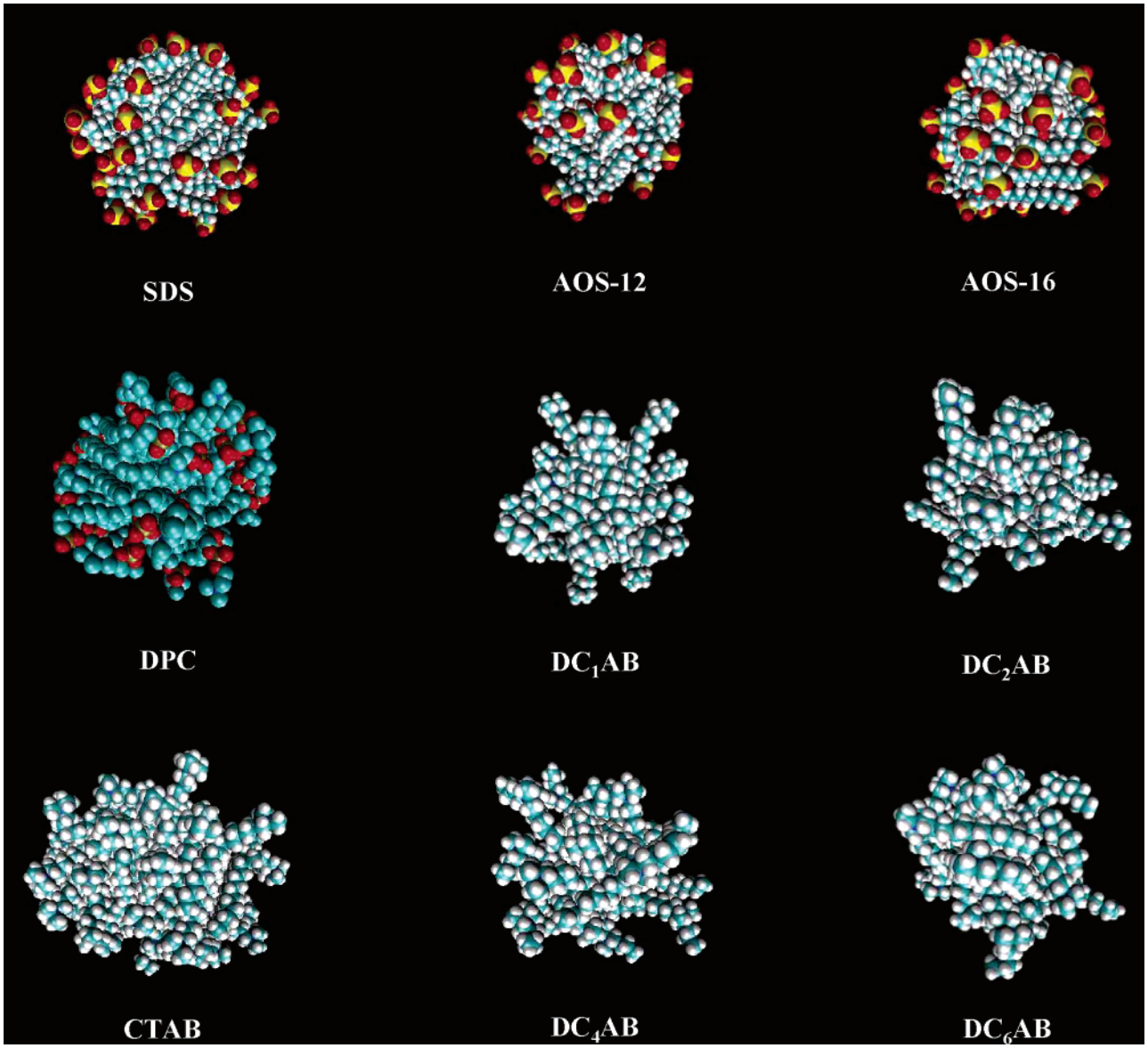


Figure 2. Snapshots of the post-equilibration structures of the simulated micelles corresponding to each of the nine ionic and zwitterionic surfactants considered here. The water molecules are not shown for clarity.

surfactants are useful because of their wetting and detergency attributes and because of their tolerance for hard water ions. AOS, as it is used industrially, is a mixture of several chemical species, including 60−70% sodium alkenesulfonate, 30% hydroxyalkanesulfonate, and 0−10% sodium disulfonate. The hydroxyalkanesulfonate fraction is present in both the 3-hydroxy sulfonate and the 4-hydroxy sulfonate forms, and the hydroxyalkanesulfonate backbone may contain between 12 and 18 carbon atoms.[23] We have selected two hydroxyalkanesulfonates for modeling to determine if computer simulations can be used to correctly identify the "head" and "tail" portions of surfactants with two hydrophilic groups (SO_3^- and OH) connected by hydrophobic CH_2 groups: sodium 3-hydroxydodecyl-1-sulfonate (AOS-12) and sodium 3-hydroxyhexadecyl-1-sulfonate (AOS-16).

The fractional degree of hydration, *f,* of AOS-12 and AOS-16 is plotted as a function of group number in Figure 6A,B, respectively. The non-monotonic nature of the fractional hydration plots of both surfactants is due to the presence of the 3-hydroxy group (group 5). The *f* results shown in Figure 6 reveal that groups 1−3 and 5 in both surfactants have *f* values greater than 0.60. The average value of *f* for these four groups in AOS-12 and AOS-16 is 0.79 and 0.75, respectively. With an *f* value of 0.60 and 0.56 in AOS-12 and AOS-16, respectively, group 4 is considered to be part of the micelle hydrophobic core for both surfactants. As such, group 4 is modeled as contributing to g_{hydr} in the CS−MT modeling approach (for details, see article 2). The average *f* value of the tail groups in AOS-12 and in AOS-16 are 0.27 and 0.21, respectively.

3.5. $C_{12}H_{25}C_NH_{2N+1}N(CH_3)_2Br$ Surfactants (DC_1AB, DC_2AB, DC_4AB, and DC_6AB). The micellization behavior of the DC_NAB cationic surfactants is very interesting from a theoretical perspective. The experimental CMCs of each of the four DC_NAB surfactants are ranked as follows: DC_1AB > DC_2AB > DC_4AB > DC_6AB.[24] This CMC ranking is difficult to rationalize using the traditional MT model (see discussion in Section 4.3.2). Because of the structural complexity of the DC_NAB surfactants, the CS−MT model is expected to provide

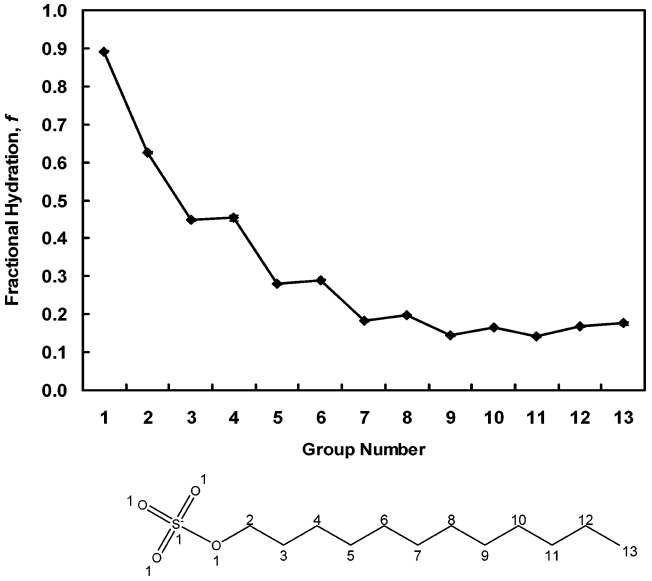


Figure 3. The average fractional degree of hydration, *f*, as defined in eq 4, of each of the groups in sodium dodecyl sulfate (SDS). The chemical structure of each group is identified in the schematic of the molecule shown below the fractional hydration plot. The error bars shown correspond to the standard error of the mean.

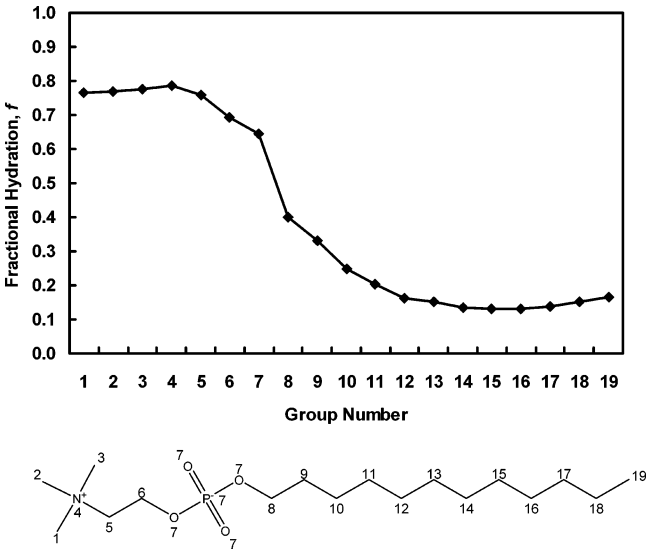


Figure 4. The average fractional degree of hydration, *f*, as defined in eq 4, of each of the groups in dodecylphosphocholine (DPC). The chemical structure of each group is identified in the schematic of the molecule shown below the fractional hydration plot. The error bars shown correspond to the standard error of the mean.

more accurate predictions of the micellization behavior of these surfactants than the traditional MT modeling approach.

The fractional degree of hydration, *f*, of DC_1AB, DC_2AB, DC_4AB, and DC_6AB is plotted as a function of group number in Figure 7A−D, respectively. The non-monotonic nature of the fractional hydration plots of the DC_2AB, DC_4AB, and DC_6AB surfactants reflects the fact that the C_N group in each of these surfactants (group 1 in DC_1AB, groups 1−2 in DC_2-AB, groups 1−4 in DC_4AB, and groups 1−6 in DC_6AB) is less hydrated than the dimethylammonium group.

The *f* values of DC_1AB shown in Figure 7A are very similar to the *f* values shown in Figure 5 for CTAB. This is not surprising given the chemical similarity of these two surfactants. The average *f* value of the groups in the DC_1AB head (groups 1−5) is 0.84. By comparison, the average *f* value of the groups in the CTAB head (groups 1−5) is 0.85.

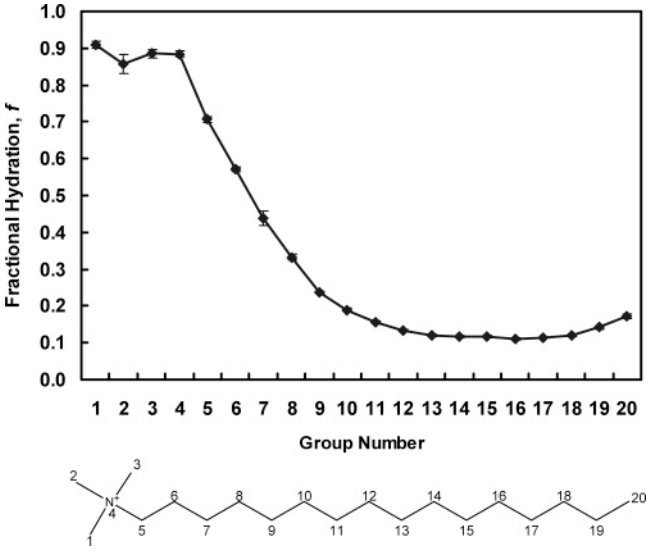


Figure 5. The average fractional degree of hydration, *f*, as defined in eq 4, of each of the groups in cetyltrimethylammonium bromide (CTAB). The chemical structure of each group is identified in the schematic of the molecule shown below the fractional hydration plot. The error bars shown correspond to the standard error of the mean.

The fractional hydration profiles of DC_2AB, DC_4AB, and DC_6AB (shown in Figure 7B−D, respectively) are more complex. Comparison of the hydration profiles of DC_1AB and DC_2AB reveals that the additional CH_2 in the C_N group of DC_2AB has a significant impact on hydration. The *f* value of group 1 in DC_2AB (0.76) is significantly lower than the *f* value of group 1 in DC_1AB (0.87). The *f* value of group 2 in DC_2AB (0.78) is also relatively low compared to the average *f* value of the groups in the DC_1AB head (0.84). The *f* values of groups 1−4 in DC_4AB range 0.71 to 0.79. The *f* values of groups 1−6 in DC_6AB range 0.57 to 0.92. Our results suggest that, as the length of the C_N group increases, the C_N group is increasingly incorporated within the micelle hydrophobic core. Groups 1 and 2 in DC_6AB both have *f* values that are below 0.60. Consequently, both of these groups will be modeled as being part of the micelle hydrophobic core using the CS−MT model.

It is interesting to note that the D group (the $C_{12}H_{25}$ group) present in each of the four DC_NAB surfactants is hydrated to a similar extent in each of the four simulated micelles. The average *f* value of the D group in DC_1AB, DC_2AB, DC_4AB, and DC_6AB is 0.34, 0.36, 0.36, and 0.31, respectively.

4. Molecular-Thermodynamic Modeling Based on Computer Simulation Inputs

4.1. Using the CS−MT Modeling Approach to Predict the Micellization Behavior of Ionic and Zwitterionic Surfactants. As discussed in article 2, to use the CS−MT modeling approach to model surfactant micellization, a number of simplifying approximations must be made to enable the straightforward evaluation of g_{dehydr} and g_{hydr}. With this in mind, we make the same approximations discussed in article 2 to evaluate g_{dehydr} and g_{hydr} for ionic and zwitterionic surfactants. In so doing, we will evaluate the validity of these approximations in the case of ionic and zwitterionic surfactants. In this article, the CS−MT model is implemented to model ionic and zwitterionic surfactant micellization in pure aqueous solution (without any added salt). The CS−MT model is also capable of modeling micellization in the presence of added salt. However, because

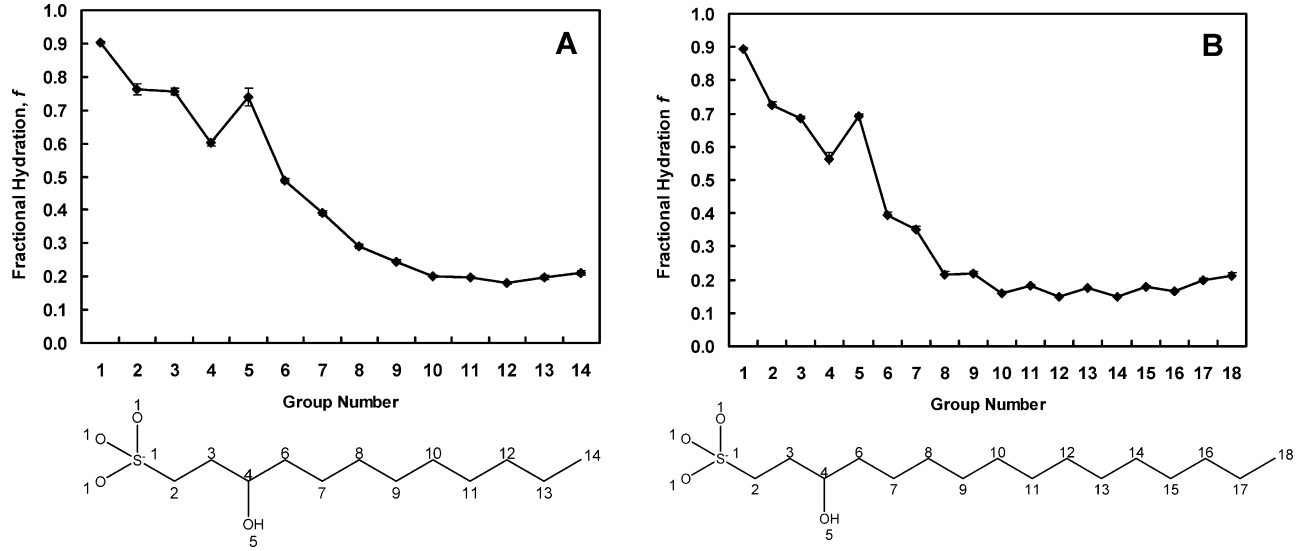


Figure 6. The average fractional degree of hydration, f, as defined in eq 4, of each of the groups in 3-hydroxy dodecyl sulfonate (AOS-12, see A) and 3-hydroxy hexadecyl sulfonate (AOS-16, see B). The chemical structure of each group is identified in the schematic of the molecule shown below each fractional hydration plot. The error bars shown correspond to the standard error of the mean.

the strength of electrostatic interactions is decreased by the presence of salt, evaluation of the approximations discussed in article 2 by modeling surfactant micellization in aqueous solution with no added salt provides the most stringent test of the applicability of these approximations for modeling ionic and zwitterionic surfactants.

After using the CS−MT model to predict the micellization behavior of two simple ionic and one simple zwitterionic surfactants in Section 4.2, we use it to model more complex surfactants in Section 4.3. To evaluate the CS−MT model, the surfactant CMC was selected for prediction and comparison with the experimental CMC data, as well as with the CMC predictions of the traditional MT model, because the CMC is exponentially dependent on g_{form} and, therefore, provides a rigorous quantitative test for the predictive accuracy of the CS−MT model. However, we would like to stress that the CS−MT model enables the prediction of a variety of micellar solution properties in addition to the CMC (including micelle shape, size, composition, and the degree of counterion binding).

As discussed in article 1, to implement the traditional MT model to predict the micellization behavior of ionic and zwitterionic surfactants, it is necessary to identify a head and a tail for each surfactant. The head and tail assignments made for traditional MT modeling are reported in Table 2, where the group numbers listed correspond to the group numbers shown in the surfactant diagrams given in Figures 3−7. For reasons that will be discussed in Section 4.3.2, the head and tail assignments for DC_2AB, DC_4AB, and DC_6AB may be made in a number of ways in the traditional MT modeling approach. On the basis of the head and tail assignments, three geometric parameters were estimated for each surfactant and used as inputs for traditional MT modeling.[3,4,25−27] These geometric parameters are also reported in Table 2 and were determined based on the surfactant chemical structures and the head and tail assignments made. The first geometric parameter is a_h—the cross-sectional area of the surfactant head. The second geometric parameter is d_{charge}—the distance from the location of the charge in the surfactant head to the beginning of the surfactant tail. The third geometric parameter is l_{hg}—the length of the surfactant head group, or the distance from the tip of the surfactant head to the start of the surfactant tail. Note that a_h is needed to calculate g_{st}, and both d_{charge} and l_{hg} are needed to calculate g_{elec}.[3,25] The a_h, d_{charge}, and l_{hg} values listed in Table 2 were also used to determine g_{st} and g_{elec} in the CS−MT model (see eq 1). Values of a_h, d_{charge}, and l_{hg} computed based on three traditional MT modeling limits (see Section 4.3.2) are also listed in Table 2 for DC_2AB, DC_4AB, and DC_6AB. An additional parameter that is not listed in Table 2 but that is needed to compute g_{elec} in the case of zwitterionic surfactants is d_{sep}—the distance between the two charges in the zwitterionic surfactant head. The d_{sep} value for DPC was estimated to be 4.3 Å.

4.2. Modeling Results for Simple Surfactants. *4.2.1. Sodium Dodecyl Sulfate (SDS).* We have used the CS−MT model to predict the micellization behavior of SDS in aqueous solution at 25 °C. In Table 3, we report CS−MT modeling results for the *simulated* SDS micelle, including (i) g_{dehydr}, (ii) g_{hydr}, (iii) $\hat{g}_{int}$, and (iv) $g_{tr,CS-MT}$. The reported uncertainty for the CS−MT modeling results is the standard error of the mean, as computed through block averaging of the computer simulation data. CS−MT modeling results for g_{dehydr} and g_{hydr} were obtained using the simulated f values and eqs 2 and 3. As can be seen, the value of g_{dehydr} ($-15.04k_BT$) is much larger in magnitude than that of g_{hydr} ($1.73k_BT$). In Table 3, we also report the traditional MT model prediction of g_{tr} for comparison with $g_{tr,CS-MT}$. We note that the transfer free-energy contribution predicted by the CS−MT model ($g_{tr,CS-MT} = -17.69k_BT$) is $0.77k_BT$ more positive than the prediction of the traditional MT model ($g_{tr} = -18.46k_BT$).

In Table 4, we report CS−MT and traditional MT modeling results for micelles of the *optimal* shape and size. As discussed in Section 1, at the optimal micelle shape and size, g_{form} attains a minimum value.[3] Both the CS−MT model and the traditional MT model yield identical predictions for the optimal micelle shape and size. As discussed in article 2, this equivalence arises because the only contribution to g_{form} that differs in the two models (the transfer free-energy contribution) does not depend on micelle shape and size. The simulated SDS micelle had an aggregation number of 44, but both the CS−MT model and the traditional MT model predict that the optimal SDS micelle is spherical with a number-average aggregation number of 47. In Table 4, we report predictions using the CS−MT model and the traditional MT model of (i) the micelle shape, (ii) the number-average micelle aggregation number (n), (iii) g_{int}, (iv) g_{pack}, (iv) g_{st}, (v) g_{elec} (including the discharging, the charging,

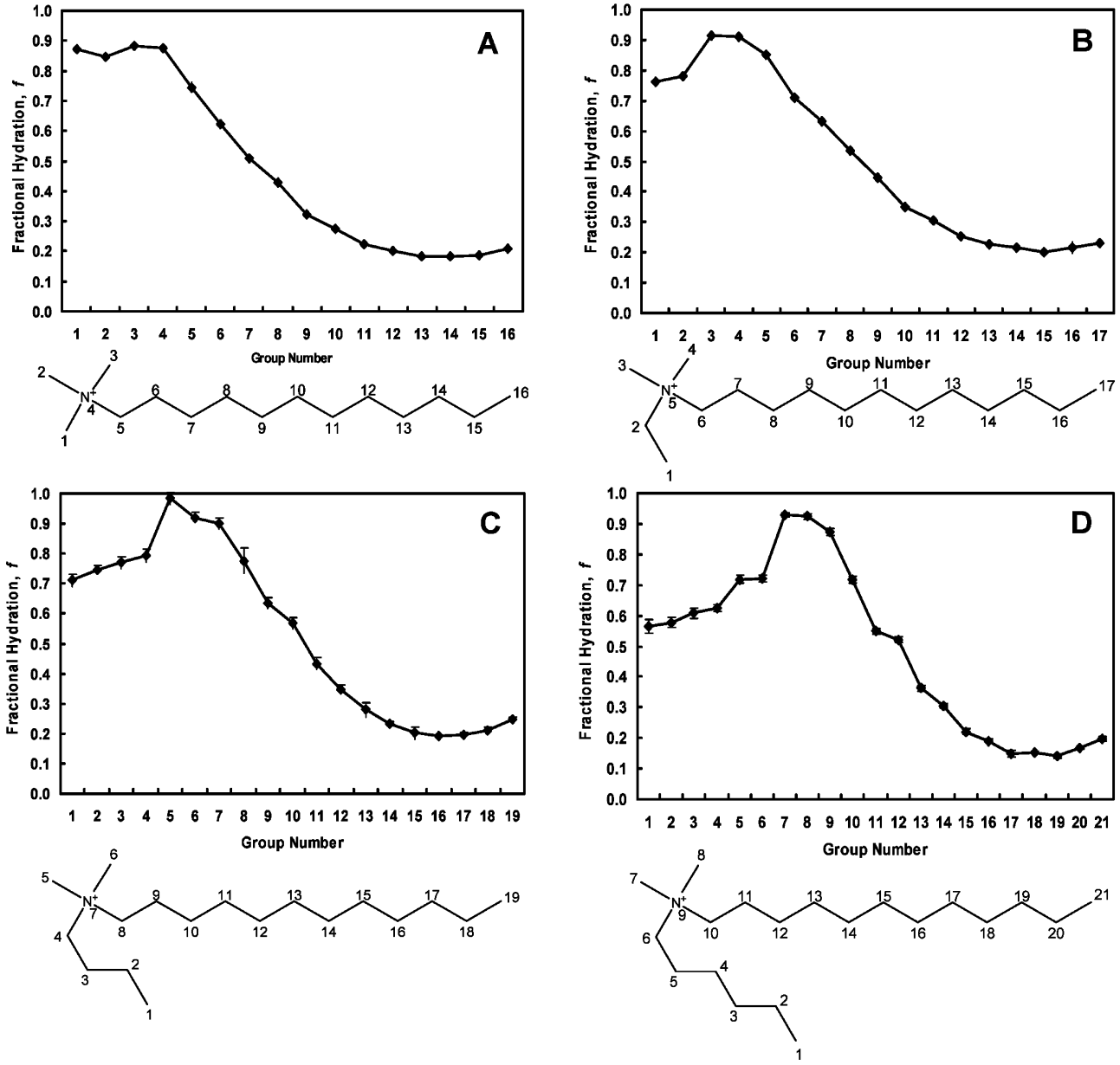


Figure 7. The average fractional degree of hydration, *f*, as defined in eq 4, of each of the groups in decyltrimethylammonium bromide (DC_1AB, see A), $C_{12}H_{25}C_2H_5N(CH_3)_2Br$ (DC_2AB,see B), $C_{12}H_{25}C_4H_9N(CH_3)_2Br$ (DC_4AB, see C), and $C_{12}H_{25}C_6H_{13}N(CH_3)_2Br$ (DC_6AB, see D). The chemical structure of each group is identified in the schematic of the molecule shown below each fractional hydration plot. The error bars shown correspond to the standard error of the mean.

and the counterion binding free-energy contributions,[25,28] (vi), g_{ent}, (vii) g_{form} and the CMC predicted by the CS−MT model, (viii) g_{form} and the CMC predicted by the traditional MT model, and (iv) the experimental values of g_{form} and of the CMC.[16] The reported uncertainty in the CS−MT modeling results corresponds to the standard error of the mean. Note that, because the shape and size of the optimal micelles predicted by the CS−MT model and the traditional MT model are identical, free-energy contributions (iii) to (vi) are also identical. Traditional MT modeling results were generated using the approach reviewed in article 1. The CMCs predicted by the CS−MT model and the traditional MT model and the value of g_{form} inferred using the experimental CMC data were calculated using eq 8.

For SDS, the value of g_{int} computed for the optimal micelle ($4.33k_BT$) is slightly lower than the value of $\hat{g}_{int}$ computed for the simulated micelle ($4.38k_BT$) because of the difference between the simulated and optimal micelle aggregation numbers. The free-energy contributions, g_{pack} ($2.51k_BT$), g_{st} ($1.10k_BT$), g_{elec} ($-0.45k_BT$), and g_{ent} ($-0.99k_BT$), although all smaller in magnitude than g_{int}, each contribute significantly to g_{form}. For this surfactant, the CMC predicted by the CS−MT model is 15.28 mM, which is roughly a factor of 2 larger than the experimental CMC (8.1 mM). The traditional MT model predicts a CMC of 9.62 mM, which is very close to the experimental CMC value of 8.1 mM.[16] Given the exponential dependence of the CMC on g_{form} (see eq 8), we consider both the CS−MT and the traditional MT results shown in Table 4 to be in reasonable agreement with the experimental data.

4.2.2. Dodecylphosphocholine (DPC). CS−MT and traditional MT modeling results for the simulated DPC micelle are reported in Table 3. Each free-energy contribution has been calculated as described in Section 4.2.1. Theoretical predictions for the optimal micelles obtained using the CS−MT model and the traditional MT model as well as experimental data[21] for the

TABLE 2: Molecular Parameters Used in Modeling Each Surfactant[a]

surfactant	head groups	tail groups	a_h [Å]^2	d_{charge} [Å]	l_{hg} [Å]
SDS	1−2	3−13	25	3.7	6.3
DPC	1−8	9−19	32	5.3	4.3
CTAB	1−5	6−20	32	3.8	6.4
AOS-12	1−5	6−14	23	5.3	7.2
AOS-16	1−5	6−18	23	5.3	7.2
DC_1AB	1−5	6−16	32	3.8	6.4
DC_2AB					
limit 1	1−6	7−17	42	3.8	6.4
limit 2	1−6	7−17	37.4	3.8	6.4
limit 3	2−6	1 and 7−17	32	3.8	6.4
DC_4AB					
limit 1	1−8	9-19	62	3.8	6.4
limit 2	1−8	9-19	38.3	3.8	6.4
limit 3	4−8	1−3 and 9−19	32	3.8	6.4
DC_6AB					
limit 1	1−10	11−21	82	3.8	6.4
limit 2	1−10	11−21	39.3	3.8	6.4
limit 3	6−10	1−5 and 10−21	32	3.8	6.4

[a] The traditional MT model identifications of heads and tails are reported, along with molecular parameters used to model each surfactant using the CS−MT and the traditional MT models. These molecular parameters were estimated geometrically on the basis of the structure of each surfactant molecule and include a_h, the cross-sectional area of the surfactant head, d_{charge}, the distance between the beginning of the surfactant tail and the location of the charge in the surfactant head, and l_{hg}, the length of the surfactant head.

micellization behavior of DPC in aqueous solution at 25 °C are reported in Table 4.

The optimal DPC micelles that are predicted to form in solution by the CS−MT model and the traditional MT model are somewhat smaller ($n = 39$) than the simulated DPC micelle ($n = 48$). The predicted value of g_{int} ($4.61k_BT$) is slightly larger than that of $\hat{g}_{int}$ ($4.26k_BT$) because of this difference in aggregation numbers. Note that g_{ent} for this surfactant is equal to zero because there is no counterion binding for this zwitterionic surfactant. The CS−MT model prediction of the transfer free-energy contribution ($g_{tr,CS-MT} = -18.32k_BT$) is slightly less negative than that of the traditional MT prediction ($g_{tr} = -18.46k_BT$). The traditional MT estimate of g_{tr} for DPC is identical to the g_{tr} estimate for SDS because both surfactants have identical tails. In contrast, the CS−MT model estimate of $g_{tr,CS-MT}$ for DPC is $0.63k_BT$ more negative than the estimate of $g_{tr,CS-MT}$ for SDS. As discussed in article 2, the nature of the surfactant head can have a significant effect on the degree of hydration of the surfactant tail in bulk aqueous solution, and this is further confirmed by the difference in the $g_{tr,CS-MT}$ values obtained for DPC and SDS. The difference between the CS−MT estimate of $g_{tr,CS-MT}$ and the traditional MT estimate of g_{tr} for DPC leads to the CS−MT model predicting a higher CMC (0.83 mM) than that predicted by the traditional MT model (0.72 mM). In this case, the CMC predicted by the CS−MT model is closer to the experimental CMC (1.0 mM).[21]

4.2.3. Cetyltrimethylammonium Bromide (CTAB). CS−MT and traditional MT modeling results for the simulated CTAB micelle are reported in Table 3. Theoretical predictions for the optimal micelles obtained using the CS−MT and traditional MT models as well as as experimental data[29] for the micellization behavior of CTAB in aqueous solution at 25 °C are reported in Table 4.

Computer simulation of CTAB was conducted in a micelle with an aggregation number of 49, which is identical (after rounding to the nearest integer value) to the CS−MT and the traditional MT number-average aggregation number predictions for the optimal CTAB micelles. The predicted value of g_{int} ($5.35k_BT$) is very close to the predicted value of $\hat{g}_{int}$ ($5.44k_BT$), and the transfer free-energy contribution ($g_{tr,CS-MT} = -23.73k_BT$) predicted by the CS−MT model is slightly less negative than that predicted by the traditional MT model ($g_{tr} = -24.43k_BT$). The difference between the CS−MT estimate of $g_{tr,CS-MT}$ and the traditional MT estimate of g_{tr} for CTAB leads to the CS−MT model predicting a higher CMC (0.81 mM) than that predicted by the traditional MT model (0.54 mM). In this case, the CMC predicted by the CS−MT model is closer to the experimental CMC (0.9 mM).[29]

4.2.4. Applicability of the CS−MT Modeling Approach to Ionic and Zwitterionic Surfactants. Based on the modeling results obtained for the simple anionic, zwitterionic, and cationic surfactants discussed in Sections 4.2.1−4.2.3, it is possible to evaluate the applicability of the CS−MT modeling approach to model the micellization behavior of surfactants with charged hydrophilic heads. For the three surfactants considered, the CMCs predicted by the CS−MT model are in reasonable agreement with the experimental CMCs. For DPC and CTAB, the CMCs predicted using the CS−MT model are more accurate than those predicted using the traditional MT model. Based on these results, we conclude that the approximations introduced in article 2 to enable straightforward calculation of g_{dehydr} and g_{hydr} are reasonably accurate for non-polymeric, small-head ionic and zwitterionic surfactants.

For the three simple, small-head surfactants modeled above, three key approximations were made to enable straightforward implementation of the CS−MT model. The first approximation involves estimating g_{tr} for hydrophobic groups in the surfactant head and tail using solubility data for linear alkyl chains. This approximation has already been shown in article 2 to yield reasonable modeling results for small-head nonionic surfactants, and we believe that it should also be reasonably accurate (and physically realistic) in the case of small-head ionic and zwitterionic surfactants. The second approximation involves identifying hydrophobic groups in each surfactant as being adsorbed onto, or incorporated within, the micelle hydrophobic core if they have an f value that is less than 0.60. It is reasonable to make this approximation for ionic and zwitterionic surfactants because the selection of this f value in article 2 was informed by data for both nonionic and ionic surfactants. The third approximation involves using the expression for Δg_{wc} given in eqs 5 and 6 to evaluate g_{hydr}. This model for Δg_{wc} was originally developed in article 1 for oil aggregates. In using such a model in the case of ionic and zwitterionic surfactant micelles, the approximation is made that the change in hydration free energy experienced by hydrophobic groups in being transferred from the bulk aqueous solution to the micelle hydrophobic core is unaffected by the presence of the charged surfactant heads and the charged counterions (if present) at the micelle core−water interface. It was not clear *a priori* whether Δg_{wc} could be evaluated for ionic and zwitterionic surfactants using eqs 5 and 6 or whether it would be necessary to fit a value of Δg_{wc} to obtain accurate predictions of the micellization behavior. Fortunately, the results for SDS, DPC, and CTAB suggest that evaluating Δg_{wc} using a model developed for oil aggregates is reasonably accurate even for ionic and zwitterionic surfactant micelles. As a result, when modeling each of the complex surfactants considered next in Section 4.3, we will calculate Δg_{wc} using eqs 5 and 6.

4.3. Modeling Results for Complex Surfactants. *4.3.1. Sodium 3-Hydroxy Sulfonates (AOS-12 and AOS-16).* CS−MT and traditional MT modeling results for the simulated AOS-12

TABLE 3: Modeling Results for the Simulated Micelles[a]

surfactant	g_{dehydr} [k_BT]	g_{hydr} [k_BT]	$\hat{g}_{int}$ [k_BT]	$g_{tr,CS-MT}$ [k_BT]	g_{tr} [k_BT]
SDS	-15.04 ± 0.03	1.73 ± 0.01	4.38	-17.69 ± 0.04	-18.46
DPC	-15.98 ± 0.34	1.93 ± 0.04	4.26	-18.32 ± 0.34	-18.46
CTAB	-20.67 ± 0.20	2.39 ± 0.04	5.44	-23.73 ± 0.21	-24.43
AOS-12	-12.74 ± 0.01	1.60 ± 0.05	4.30	-15.44 ± 0.05	-15.37
AOS-16	-18.50 ± 0.10	1.95 ± 0.03	5.26	-21.81 ± 0.10	-21.27
DC_1AB	-14.80 ± 0.07	1.44 ± 0.03	5.10	-18.46 ± 0.07	-18.46
DC_2AB	-14.96 ± 0.10	1.44 ± 0.04	5.54	-19.06 ± 0.11	-18.46 to -21.98
DC_4AB	-15.44 ± 0.10	1.10 ± 0.03	5.87	-20.21 ± 0.11	-18.46 to -24.96
DC_6AB	-18.41 ± 0.11	2.71 ± 0.07	6.55	-22.24 ± 0.13	-18.46 to -27.95

[a] CS−MT and traditional MT modeling results for each of the nine simulated ionic and zwitterionic surfactant micelles considered in this article. CS−MT model predictions of g_{dehydr}, g_{hydr}, $\hat{g}_{int}$, and $g_{tr,CS-MT}$ were made as described in Section 1.1. The uncertainties reported for the CS−MT model predictions correspond to the standard error of the mean. Traditional MT modeling results for g_{tr} are presented for comparison with $g_{tr,CS-MT}$.

TABLE 4: Modeling Results for the Optimal Micelles[a]

							g_{form} [k_BT] (CMC [mM])		
surfactant	shape	n	g_{int} [k_BT]	g_{pack} [k_BT]	g_{st} [k_BT]	g_{elec} [k_BT]	g_{ent} [k_BT]	CS-MT model	traditional MT model	exptl
SDS	sph	47	4.33	2.51	1.10	3.56	−0.99	-8.20 ± 0.04 (15.28 ± 0.58)	−8.66 (9.62)	−8.83 (8.1)
DPC	sph	39	4.61	2.47	0.65	0.48	0.00	-11.11 ± 0.34 (0.83 ± 0.29)	−11.25 (0.72)	−10.92 (1.0)
CTAB	sph	49	5.35	2.57	1.07	5.54	−0.93	-11.14 ± 0.21 (0.81 ± 0.17)	−11.55 (0.54)	−11.03 (0.9)
AOS-12	sph	21	4.95	1.28	0.61	2.59	−0.71	-7.87 ± 0.05 (21.29 ± 1.02)	−7.83 (22.18)	−7.71 (24.8)
AOS-16	sph	45	4.83	2.53	0.77	4.82	−0.87	-10.74 ± 0.10 (1.21 ± 0.13)	−10.42 (1.66)	−10.55 (1.45)
DC_1AB	sph	47	4.24	2.87	1.45	3.67	−0.97	-8.18 ± 0.07 (15.51 ± 1.13)	−8.22 (12.59)	−8.16 (15.90)
DC_2AB	sph	47	4.24	2.83	1.78	3.85	−0.93	-8.29 ± 0.11 (13.99 ± 1.52)	−9.85 to −7.65 (2.93 to 26.42)	−8.23 (14.80)
DC_4AB	sph	47	4.24	2.85	1.82	4.30	−0.92	-8.92 ± 0.11 (7.44 ± 0.80)	−11.56 to −6.33 (0.53 to 51.66)	−8.82 (8.20)
DC_6AB	sph	41	4.49	2.92	1.67	4.97	−0.88	-10.08 ± 0.13 (2.33 ± 0.30)	−13.73 to −5.06 (0.06 to 354.08)	−9.59 (3.80)

[a] CS−MT and traditional MT modeling results for each of the nine ionic and zwitterionic surfactants considered in this article. Both the CS−MT model and the traditional MT model yield identical predictions of the optimal micelle shape, the number-average micelle aggregation number (n), g_{int}, g_{pack}, and g_{st} (see Section 4.2.1). The CS−MT and the traditional MT model predictions of g_{form} were obtained using the values of $g_{tr,CS-MT}$ and g_{tr} reported in Table 4, respectively, as an input to eq 7. The CS−MT and the traditional MT model predictions of the CMC and the value of g_{form} inferred from the experimental CMC data were computed using eq 8. The uncertainties reported for the CS−MT model predictions correspond to the standard error of the mean.

and AOS-16 micelles are reported in Table 3. Theoretical predictions for the optimal micelles obtained using the CS−MT model and the traditional MT model as well as experimental data[30] for the micellization behavior of AOS-12 and AOS-16 in aqueous solution at 30 °C are reported in Table 4.

The CS−MT and the traditional MT model predictions of the optimal number-average micelle aggregation numbers of AOS-12 ($n = 21$) and AOS-16 ($n = 45$) are somewhat different than the aggregation numbers of the simulated AOS-12 ($n = 32$) and AOS-16 ($n = 40$) micelles. The predicted value of g_{int} for AOS-12 ($4.95k_BT$) is significantly larger than that of $\hat{g}_{int}$ ($4.30k_BT$) because of this difference in aggregation numbers. In contrast, g_{int} for AOS-16 ($4.83k_BT$) is predicted to be significantly smaller than $\hat{g}_{int}$ ($5.26k_BT$). For both AOS-12 and AOS-16, the CS−MT model predictions of the transfer free-energy contribution ($g_{tr,CS-MT} = -15.44$ and $-21.81k_BT$, respectively) are slightly more negative than the traditional MT model predictions of the transfer free-energy contribution ($g_{tr} = -15.37$ and -21.27, respectively). This leads to the CS−MT model predicting lower CMCs for both surfactants (21.29 and 1.21 mM for AOS-12 and AOS-16, respectively) than those predicted by the traditional MT model (22.18 and 1.66 mM, respectively). The CMCs predicted by the CS−MT and the traditional MT models are both in reasonable agreement with the experimental CMCs (24.8 and 1.45 mM, respectively). However, it is important to note that the groups in AOS-12 and AOS-16 that should be modeled as being part of the surfactant head and as part of the surfactant tail in traditional MT modeling are not entirely clear. As a result, we have used head and tail assignments for these surfactants determined through computer simulation in a previous study.[26] Without such information, it would not have been possible to make such accurate predictions using the traditional MT model.

For these two surfactants, our results indicate that, although computer simulation inputs are necessary, reasonably accurate predictions of their micellization behavior can be made using computer simulations to make head and tail identifications for use in traditional MT modeling (as was done in ref 26) or to obtain fractional hydration data for use in the CS−MT model (as was done in this article). In ref 26, head and tail identifications for both of these surfactants were made in a computationally efficient way by simulating both surfactants at a water−oil interface (serving as a proxy for the micelle core−water interface), and comparably accurate modeling results were obtained. The four computer simulations used here to obtain fractional hydration information for these two surfactants for input in the CS−MT model required approximately 1 order of magnitude more computational expense.

4.3.2. $C_{12}H_{25}C_NH_{2N+1}N(CH_3)_2Br$ Surfactants (DC_1AB, AB, DC_2AB, DC_4AB, and DC_6AB). CS−MT and traditional MT modeling results for the simulated DC_1AB, DC_2AB, DC_4AB, and DC_6AB micelles are reported in Table 3. Theoretical predictions for the optimal micelles obtained using the CS−MT and the traditional MT models as well as experimental data[30] for the micellization behavior of each surfactant in aqueous solution at 25 °C are reported in Table 4. The approach described in Section 4.2.1 was used to calculate each free-energy contribution, the values of g_{form}, and the CMC values reported in Table 4.

As shown in Tables 3 and 4, although only one value of $g_{tr,CS-MT}$, g_{form}, and the CMC are reported for DC_2AB, DC_4AB, and DC_6AB using the CS−MT model, a range of g_{tr}, g_{form}, and CMC values are reported for these surfactants using the traditional MT modeling approach because it is difficult to determine the head and tail of each surfactant. As discussed in article 1, for simple surfactants and solubilizates, simple guidelines can be used for head and tail identification. For ionic and zwitterionic surfactants, the approximation is made that $n_t = n_c - 1$, where n_c is the total number of CH_2 and CH_3 groups in the hydrocarbon chain and n_t is the number of CH_2 and CH_3 groups that should be modeled as being part of the surfactant tail.[3,31] In other words, a hydrophobic group bonded to a charged hydrophilic group is modeled as being part of the surfactant head, and other hydrophobic groups are modeled as being part of the surfactant tail. Unfortunately, these guidelines are inadequate to provide accurate head and tail identifications for DC_2AB, DC_4AB, and DC_6AB because it is unclear how to model each C_N group. Because of this uncertainty, we have implemented the traditional MT modeling approach based on three different modeling limits. The first limit (which we will refer to as limit 1) involves modeling each atom in the C_N groups as being part of the surfactant head, with the remaining groups in each surfactant being modeled as being part of the head or the tail according to the $n_t = n_c - 1$ guideline. Using this approach, groups 1−6 of DC_2AB are modeled as being part of the head, and groups 7−17 are modeled as being part of the tail (see the group numbers defined in Figure 7 and the a_h values listed in Table 2). On the basis of the structure of the selected head, the area of the surfactant head (a_h) was computed geometrically by assuming that the entire C_N group lies parallel to the micelle core−water interface and contributes to the surfactant head area at the interface. The second limit (limit 2) is based on the same assignment of head and tail made in limit 1. However, in limit 2, an accurate value of a_h for each surfactant has been determined from the computer simulation data by measuring the projected area of each surfactant head at the micelle core−water interface. The same values of a_h used in limit 2 of traditional MT modeling were also used in CS−MT modeling. It is important to note that the estimation of a_h using the computer simulation data is expected to yield significantly more accurate results for complex surfactants such as DC_2AB, DC_4AB, and DC_6AB than attempting to approximate a_h based on an assumption of what portions of the surfactant are present at the micelle core−water interface. In the third limit (limit 3), the head and tail portions of DC_2AB, DC_4AB, and DC_6AB were estimated using the $n_t = n_c - 1$ guideline for both the C_N chain and each of the remaining groups in the surfactant. Using this approach, groups 2−6 of DC_2AB were assigned as being part of the head, and groups 1 and 7−17 were assigned as being part of the tail (see the group numbers defined in Figure 7 and the a_h values listed in Table 2). Similarly, groups 4−8 of DC_4AB were assigned as being part of the head, and groups 1−3 and 9−19 were assigned as being part of the tail. Based on this assignment of heads and tails, the values of a_h for DC_2AB, DC_4AB, and DC_6AB are identical and equal to the value of a_h for DC_1AB. Limits 1 and 3 yield upper- and lower-bound estimates of the CMC using the traditional MT modeling approach, respectively. As demonstrated by the results shown in Table 4, the range of CMC values encompassed by these two limits is very large, and the range of values increases in magnitude as the length of the C_N group increases.

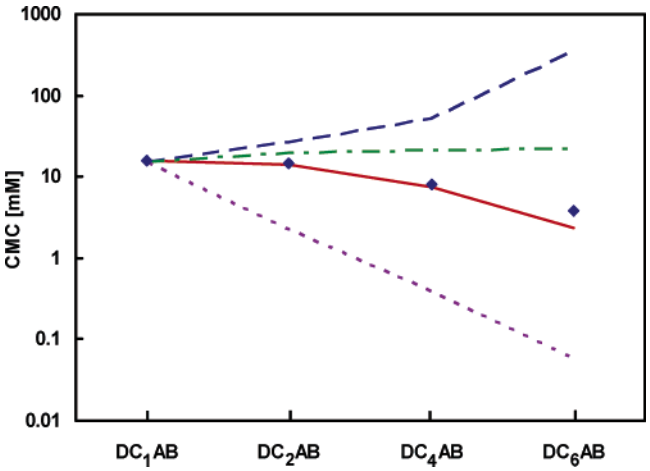


Figure 8. Comparison of the CMCs predicted using the CS−MT model (—) and the CMCs predicted using the three traditional MT modeling limits discussed in Section 4.3.2 (limit 1 — —, limit 2 — ·, and limit 3 - - -) for decyltrimethylammonium bromide (DC_1AB), $C_{12}H_{25}C_2H_5N(CH_3)_2Br$ (DC_2AB), $C_{12}H_{25}C_4H_9N(CH_3)_2Br$ (DC_4AB), and $C_{12}H_{25}C_6H_{13}N(CH_3)_2Br$ (DC_6AB). The lines are shown as a guide for the eye. The experimental CMC data is shown for comparison (◆).

The CMCs predicted using the CS−MT and the traditional MT models are reported graphically in Figure 8 on a log scale. Not only are the CMCs predicted using the CS−MT model in remarkably good agreement with the experimental CMCs, but they are also much more accurate than any of the CMCs predicted using the three traditional MT modeling limits.

The traditional MT model fails to accurately model the DC_2AB, DC_4AB, and DC_6AB surfactants for several reasons. Limit 1 is in poor agreement with the experimental CMC data because (i) computing a_h geometrically based on the assumption that the entire C_N group is part of the surfactant head and lies flat at the micelle core−water interface overestimates a_h and consequently leads to an overestimation of g_{st}, and (ii) making the approximation that the C_N group is part of the surfactant head (and therefore remains fully hydrated upon micelle formation) leads to a severe overprediction of g_{tr}. Limit 2 is in poor agreement with the experimental CMC data because of (ii). It is interesting to note, however, that using a computer simulation estimate of a_h significantly improves the traditional MT model CMC predictions. Limit 3 is in poor agreement with the experimental CMC data because (i) the head and tail assignments made in this limit imply that a_h for DC_2AB, DC_4AB, and DC_6AB are each equal to 32 Å^2, which underestimates the value of a_h, and (ii) the large number of groups included in the surfactant tail leads to an overly negative estimate of g_{tr}.

The length of the C_N group in each DC_NAB surfactant influences the CS−MT model estimate of g_{form} in several interesting ways. First, as the length of the C_N group increases, the computer simulation estimate of a_h increases. This serves to increase the steric free-energy contribution, g_{st}. At the same time, however, the transfer free-energy contribution, $g_{tr,CS-MT}$, of the C_N group becomes increasingly negative. In addition, the

C_N group in DC_6AB has two hydrophobic groups (1 and 2) that are sufficiently dehydrated that they are modeled as being part of the micelle core and that therefore affect g_{pack}. As shown in Table 4, the experimental CMC values of the four DC_NAB surfactants are ranked as follows: $DC_1AB > DC_2AB > DC_4AB > DC_6AB$.[24] Consequently, the net effect of increasing the length of the C_N group is to lower g_{form}, which in turn lowers the value of the CMC. Not only is this ranking of g_{form} and CMC values correctly predicted by the CS−MT model, but the values of g_{form} and the CMC predicted by the CS−MT model are in very close agreement with the experimental data.

In article 2, we concluded that care must be taken in computing g_{dehydr} for surfactants with large, polymeric heads using the simplifying approximations discussed in Section 4.2.4. It is important to note that the DC_NAB surfactants modeled here do not fall into this category of surfactants for the following reasons: (i) the maximum size of the DC_NAB surfactant heads is relatively small, and (ii) hydrophobic groups in the DC_NAB heads are transferred (to the extent that they are dehydrated) to an environment that is chemically similar to a bulk solution of surfactant tails (see additional discussion in article 2).

Evaluation of the majority of the free-energy contributions that appear in eq 7 for the DC_NAB surfactants is relatively straightforward. However, accurate estimation of the packing free-energy contribution, g_{pack}, for limit 3 of the traditional MT and CS−MT modeling of DC_2AB, DC_4AB, and DC_6AB has required the development and implementation of a modified packing model that is capable of accurately estimating g_{pack} for surfactants with two tails separated by a head. The packing free-energy contribution, g_{pack}, represents the free-energy change required to fix one end of the surfactant tail(s) at the micelle core−water interface. This free-energy contribution is typically estimated using a mean-field model first introduced by Ben-Shaul, Szleifer, and Gelbart[32−34] and requires sampling each important conformation and orientation of the surfactant tail subject to the constraint that the micelle hydrophobic core has a uniform density. As shown in Table 3, in limit 3 of the traditional MT modeling of DC_2AB, tail groups 1 and 7−17 are separated by head groups 2−6. A similar separation of head and tail groups is also present in DC_4AB and DC_6AB. To maintain consistency between the CS−MT model for the hydrophobic effect (g_{dehydr} and g_{hydr}) and the evaluation of g_{pack} for the CS−MT model, head and tail assignments used in computing g_{pack} for the CS−MT model were made using the criterion that any group with an f value greater than 0.60 was modeled as a group in the head, and any group with an f value less than 0.60 was modeled as a group in the tail. Using this approach, each of the C_N groups in DC_2AB and DC_4AB was identified as being part of the head. Groups 3−6 in the C_N group in DC_6AB were identified as being part of the head, and groups 1−2 were identified as being part of the tail.

The mean-field approach that was implemented to determine g_{pack} for limit 3 of the traditional MT model and the CS−MT model was modified to be more physically realistic by relaxing the constraint that groups in the surfactant tail cannot exit the micelle core. In the modified approach, each important conformation and orientation of the C_N group was sampled, and the C_N group was modeled as contributing to the volume of the micelle core for conformations where atoms in the C_N group entered the micelle core. For conformations where atoms in the C_N group did enter the micelle core, they affect the lateral pressures present in the micelle core and the value of g_{pack}. This packing approach allowed sampling of each of the conformations that were actually observed during the molecular dynamics simulation, where each C_N group was found to adopt a wide variety of conformations both inside and outside the micelle core. Complete details of the modified packing approach used to compute g_{pack} for DC_2AB, DC_4AB, and DC_6AB will be presented in a future publication.

Although not explored in this article, the predictions of the traditional MT model may be improved by using computer simulations to determine the appropriate head and tail assignments for the DC_2AB, DC_4AB, and DC_6AB surfactants. Such an approach (with data taken from a previous publication)[26] was used in modeling AOS-12 and AOS-16 in Section 4.3.1. Although computer simulation determination of head and tail groups might improve the traditional MT model predictions, this approach is unlikely to yield predictions that are as accurate as those obtained using the CS−MT model. This is due to the fact that the traditional MT model is limited by the simplistic modeling approximation that groups in the surfactant head remain fully hydrated in the micellar state. The effect of this approximation can be understood by closely examining the traditional MT modeling results for DC_2AB. In limit 2, the entire C_N group (CH_2-CH_3) is modeled as being part of the head, and in limit 3, the CH_2 group in C_N is modeled as being part of the head and the CH_3 group in C_N is modeled as being part of the tail. These two limits represent the only physically plausible head and tail assignments for the C_N group in DC_2AB that could be obtained from computer simulation. As shown in Figure 8, both limits yield predicted CMCs that are less accurate than the CMC predicted using the CS−MT model.

5. Conclusions

In this article, we have demonstrated the validity and accuracy of the CS−MT model by using it to model nine ionic and zwitterionic surfactants of varying structural complexity. To implement the CS−MT model, we have used molecular dynamics computer simulations to determine quantitative information about the changes in hydration that occur upon micelle formation. This detailed hydration information was then used to quantify the hydrophobic driving force for micelle self-assembly ($g_{tr,CS-MT}$). After determining this input, the free energy of micelle formation, g_{form}, and the CMC were calculated for each surfactant for micelles of the optimal shape and size.

To quantify the hydration changes that occur upon micelle formation, we conducted two independent molecular dynamics simulations for each of the nine ionic and zwitterionic surfactants modeled. Changes in hydration were quantified by computing a fractional hydration value, f, for each group. The f values obtained for each surfactant through MD simulation were used as an input in a free-energy model to compute the magnitude of the hydrophobic driving force for micelle formation (g_{dehydr} and g_{hydr}). In this article, we have used the approximations discussed in article 2 to estimate g_{dehydr} and g_{hydr}. As discussed in Section 4.2.4, three key approximations were tested, including the approximation that Δg_{wc} can be evaluated with reasonable accuracy using the same expression for nonionic and ionic/zwitterionic surfactant systems. All three approximations were found to be reasonably accurate in modeling ionic and zwitterionic surfactants. In particular, we found that Δg_{wc} could be modeled using eqs 2 and 3 and need not be fitted to obtain accurate results for ionic and zwitterionic surfactants.

Reasonable agreement between the CS−MT model predictions and the experimental data for g_{form} and the CMC were obtained for each of the nine ionic and zwitterionic surfactants modeled in this article. For six of these surfactants (SDS, DPC, CTAB, AOS-12, AOS-16, and DC_1AB), the CMCs predicted using the CS−MT model were found to be in reasonable

agreement with the CMCs predicted using the traditional MT model. However, for DC_2AB, DC_4AB, and DC_6AB, the predictions of the CS−MT model were in much closer agreement with the experimental data than the predictions of the traditional MT model.

The results obtained for the relatively complex surfactants DC_2AB, DC_4AB, and DC_6AB highlight the strengths of the CS−MT modeling approach: for surfactants where a significant number of hydrophobic groups are located near the aggregate core−water interface and remain partially hydrated upon micelle formation, the CS−MT modeling approach eliminates the guesswork involved in traditional MT modeling. Furthermore, because the CS−MT modeling approach uses a more realistic free-energy model to quantify the hydrophobic driving force for micelle formation, it yields more accurate predictions of the micellization behavior than the traditional MT modeling approach. Given the relatively small fraction of surfactants with sufficient structural and chemical simplicity to be accurately modeled using the traditional MT modeling approach, we believe that the CS−MT model represents a very attractive and useful alternative.

Acknowledgment. This research was supported in part by funding provided by DuPont through the DuPont−MIT Alliance. B.C.S. would like to thank Arthur Goldsipe for stimulating discussions about the CS−MT model. B.C.S. would also like to thank Jonathan Mendenhall for developing a mean-field modeling approach to compute g_{pack} for surfactants with complex head and tail structures, thereby enabling evaluation of g_{pack} the DC_NAB surfactants modeled in this article.

Supporting Information Available: Two figures are available that plot normalized SASA values as a function of simulation time during the 5 ns data-gathering simulation run (see Section 2.3) for each of the nine ionic and zwitterionic surfactant micelles simulated in this article. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Stephenson, B. C.; Goldsipe, A.; Beers, K. J.; Blankschtein, D. *J. Phys. Chem. B* **2007**, *111*, 1025−1044.

(2) Stephenson, B. C.; Goldsipe, A.; Beers, K. J.; Blankschtein, D. *J. Phys. Chem. B* **2007**, *111*, 1045−1062.

(3) Puvvada, S.; Blankschtein, D. *J. Chem. Phys.* **1990**, *92*, 3710−3724, and references cited therein.

(4) Goldsipe, A.; Blankschtein, D. *Langmuir* **2005**, *22*, 9850−9865, and references cited therein.

(5) Smith, R.; Tanford, C. *Proc. Natl. Acad. Sci.* **1973**, *70*, 289−293.

(6) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.

(7) Schweighofer, K. J.; Essmann, U.; Berkowitz, M. *J. Phys. Chem. B* **1997**, *101*, 3793−3799.

(8) Tieleman, D. P.; Berendsen, H. J. C. *J. Chem. Phys.* **1996**, *105*, 4871−4880.

(9) van der Spoel, D.; Lindahl, E.; Hess, B.; van Buuren, A. R.; Apol, E.; Meulenhoff, P. J.; Tieleman, D. P.; Sijbers, A. L. T. M.; Feenstra, K. A.; van Drunen, R.; Berendsen, H. J. C. *Gromacs User Manual*, 2004.

(10) Chirlian, L. E.; Francl, M. M. *J. Comput. Chem.* **1987**, *8*, 894−905.

(11) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.

(12) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577−8593.

(13) van der Spoel, D.; Lindahl, E.; Hess, B.; van Buuren, A. R.; Apol, E.; Meulenhoff, P. J.; Tieleman, D. P.; Sijbers, A. L. T. M.; Feenstra, K. A.; van Drunen, R.; Berendsen, H. J. C. *Gromacs User Manual version 3.2*; www.gromacs.org:2004.

(14) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. *Comput. Phys. Commun.* **1995**, *91*, 43−56.

(15) Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Model.* **2001**, *7*, 306−317.

(16) Israelachvili, J. N. *Intermolecular and Surface Forces*, 2nd ed.; Academic Press: New York, 1991.

(17) Bruce, C.; Berkowitz, M.; Perera, L.; Forbes, M. D. E. *J. Phys. Chem. B* **2002**, *106*, 3788−3793.

(18) Flyvbjerg, H.; Petersen, H. G. *J. Chem. Phys* **1989**, *91*, 461−466.

(19) Hess, B. Thesis. *Stochastic Concepts in Molecular Simulation*; Rijksuniversiteit Groningen, 1999.

(20) Hess, B. *J. Chem. Phys.* **2001**, *116*, 209−217.

(21) Kallick, D. A.; Tessmer, M. R.; Watts, C. R.; Li, C.-Y. *J. Magn. Reson.* **1995**, *109*, 60−65.

(22) Tieleman, D. P.; van der Spoel, D.; Berendsen, H. J. C. *J. Phys. Chem. B* **2000**, *104*, 6380−6388.

(23) Van Os, N. M.; Van Ginkel, R.; Van Zon, A.; Heywood, F. W.; Berryman, E. L.; Borchardt, J. K. *Solubilization in Surfactant Aggregates. Surfactant Science Series 56*; Marcel Dekker: New York, 1996.

(24) Bai, G. Y.; Wang, J. B.; Yan, H.; Li, Z. X.; Thomas, R. K. *J. Phys. Chem. B* **2001**, *105*, 9576−9580.

(25) Srinivasan, V.; Blankschtein, D. *Langmuir* **2003**, *19*, 9932−9945, and references cited therein.

(26) Stephenson, B. C.; Beers, K.; Blankschtein, D. *Langmuir* **2006**, *22*, 1500−1513, and references cited therein.

(27) Stephenson, B. C.; Rangel-Yagui, C. O.; Pessoa, A.; Tavares, L. C.; Beers, K.; Blankschtein, D. *Langmuir* **2006**, *22*, 1514−1525.

(28) Srinivasan, V.; Blankschtein, D. *Langmuir* **2003**, *19*, 9946−9961.

(29) Hansson, P.; Jonsson, B.; Strom, C.; Soderman, O. *J. Phys. Chem. B* **2000**, *104*, 3496−3506.

(30) Van Os, N. M.; Rupert, L. A. M.; Haak, J. R. *Physico-Chemical Properties of Selected Anionic, Cationic and Nonionic Surfactants*; Elsevier: Amsterdam, 1993.

(31) Srinivasan, V. Thesis. *Theoretical Modeling of Micellization and Solubilization in Ionic Surfactant Systems*; Massachusetts Institute of Technology, 2003, and references cited therein.

(32) Ben-Shaul, A.; Szleifer, I. *J. Chem. Phys.* **1985**, *83*, 3597−3611.

(33) Szleifer, I.; Ben-Shaul, A.; Gelbart, W. M. *J. Chem. Phys.* **1985**, *83*, 3612−3620.

(34) Szleifer, I.; Ben-Shaul, A.; Gelbart, W. M. *J. Chem. Phys.* **1987**, *86*, 7094−7109.